# PHARMACOKINETICS, CELL KINETICS AND PHARMACODYNAMICS. A BAYESIAN APPROACH.


Haritz Olaeta

School of Mathematics

August 2002

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Science

PAGE NUMBERING AS IN THE ORIGINAL THESIS

# Acknowledgements

First of all I would like to thank my supervisor Professor Peter Green for his constant support and for being patient with me. I have been fortunate to benefit not only from his expertise in statistics but also from his broad knowledge and understanding of science.

Thanks also to Dr. Stephen Lowis for teaching me everything I know about pharmacokinetics, pharmacodynamics and cell kinetics and for providing very valuable data. To Dr. Sean Collins and Professor Bernard Silverman for their encouragement and support in the yearly reviews. To the Basque Government for the financial support.

To my office-mates Yanan Fan, Su Subarao, Scott Sisson, Natalia Botchkina and Cody Cotar for their friendship and support and to all the people in the department for making it such a nice environment to work in. To my friends Augusto Mendez de Lugo, Pepe Tribo, Pablo Tejera, Bego de Torres, Ibon Galarraga, Carmen Alba, Marta Gonzalez, Patti Battelli, Ingrid Cabades, Isabel Ortola, Raimondo Cecchini, Alberto Riminucci, Esther Amaya and Dulce Munoz, thank you for making my life so enjoyable during my stay in Bristol.

Last, but certainly not least, to my parents for their unconditional love.

# Author's Declaration

I declare that the work in this thesis was carried out in accordance with the Regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and do not necessarily represent those of the University of Bristol. The thesis has not been presented to any other university for examination either in the United Kingdom or overseas.

Haritz Olaeta

Date: 1st August 2001

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Motivation and General Outline

## 1.1 Pharmacological Motivation

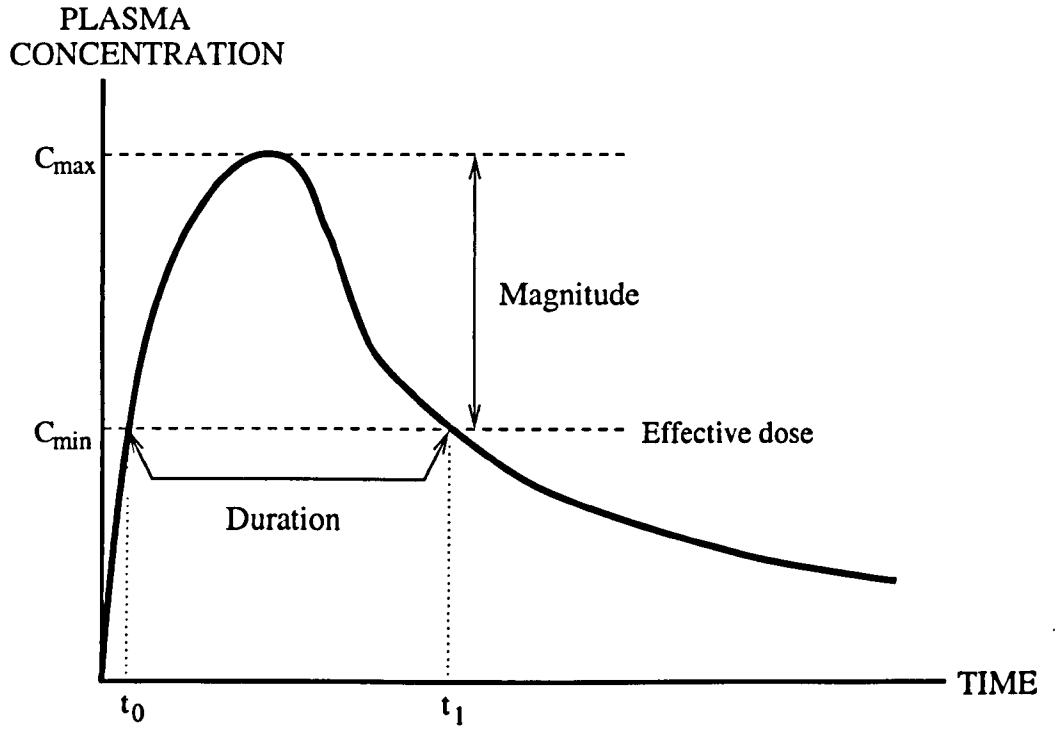
### 1.1.1 Pharmacokinetics

When studying the characteristics and rational uses of a drug, the intensity and duration of action play a vital role. These two parameters are related to the concentration of drug at its site of action and the time during which the effective amount of drug remains active there.

As the concentration of drug at its site of action is in most cases unassessable, the more accessible measure of blood plasma drug concentration is most often used to determine the intensity of drug at its site of action provided that there exists a good understanding of the processes involved in the evolution of the drug within the body. In determining drug action is, therefore, crucial to measure drug blood plasma concentration at different times and to understand the processes involved in its evolution.

Figure 1.1 illustrates the effect of a drug as a function of drug concentration in blood plasma over time. For that purpose, a plausible blood plasma drug

Figure 1.1: Drug concentration and effect. A typical blood plasma drug concentration following a single dose orally administered. The total effect of the dose will be determined by the time that drug concentration levels exceed a minimum concentration level, $C_{min}$ , required in order to produce an effect. Blood concentration levels should avoid to reach toxicity levels, $C_{max}$.




concentration–time relationship is drawn. It is shown that immediately after administration, drug concentration in blood rises rapidly as the substance is being conducted into the circulatory system. Eventually, when no more drug is to enter the circulatory system, peak plasma concentration will be reached, and it will start decreasing as the drug is being delivered into other organ and tissues and is being removed from the body.

The rates of absorption and distribution govern the time of onset of drug's action; the rates of metabolism and excretion govern its duration; while the size of the dose, in combination with these effects, governs the intensity.

The degree of biological effect produced by a drug is a function of the amount

of drug at the target site which, at the same time, is a function of the amount of drug found in the circulatory system. Only a fraction of the quantity reported in the blood plasma will reach the desired site of action. This last fraction of the dose is the only one that will actually be effective for therapeutical purposes and, consequently, the one that will determine the final intensity and duration of the aimed effect at the target organ or tissue. For instance, in the concentration-time profile presented in figure 1.1, a 'minimum concentration', $C_{\text{min}}$, has to be encountered in the blood plasma sample for the drug to reach the target site at a effective amount. For most drugs a quantity as important to determine as this minimum concentration level is the 'maximum concentration', $C_{\text{max}}$. Above this level toxicity may occur as the substance can have side effects in different organ and tissues.

A good understanding of the drug flow within the body would lead to optimal doses given that there is a knowledge of the actual effect that certain drug concentration profiles produce in different parts of the body.

Pharmacokinetics is the study of the absorption, distribution and elimination of drugs with respect to time. The quantitative importance of each of these processes for a given drug will determine the ultimate drug concentration achieved at the desired site of action and the implied level of toxicity. Gibaldi and Perrier (1982) constitutes an excellent introduction to the subject.

The discipline of pharmacokinetics, and the related discipline of drug metabolism, have evolved rapidly from being minor, almost insignificant contributors to the drug discovery and development process, to the present situation where they play a major, pivotal role in all phases from early discovery through development, and beyond. The increasing role of pharmacokinetics in drug discovery and development has emerged as a result of maturation of scientific awareness of its importance and also advances in analytical and computer technology.

By necessity, pharmacokinetics is highly interdisciplinary. These interactions reflect the ultimate purpose of pharmacokinetics, which is not merely to mathematically describe the time course of drug disposition, but to provide one of the critical scientific components that contribute to understanding of drug efficacy and safety.

### 1.1.2 Pharmacokinetics and Pharmacodynamics

It is broadly recognised that drug effect may on occasion be directly associated with the time course of drug or drug metabolites in the circulation. Nevertheless, the relationship between drug concentration and response can be quite complex such that the temporal relationship might be difficult to discern on first examination. An exhaustive quantitative description of the dose-concentration-effect-time relationship is required for a better understanding of the drug course in the patients' organism. Pharmacodynamic models relate some response measurement to the dose or to some function of the plasma drug concentration.

The mathematical modelling activity of the dose-concentration-effect-time relationship stimulates the integration of both pharmacokinetic and pharmacodyamic behaviour during both preclinical and clinical trials. Pharmacokinetic-Pharmacodynamic (PK/PD) models can lend substantial support to hypotheses of pharmacological mechanisms of action in complex biological systems that would otherwise be in most cases untractable.

Knowledge that a specific test for pharmacological activity gives rise to exposure values predictive of, but not necessarily the same as, values in humans can facilitate initial dose-ranges studies in preclinical animal trials (Collins *et al.* 1990). For example, Unadkat *et al.* (1986) developed an integrated model for the interaction of nondepolarising neuromuscular blocking drugs and their

antagonists based on canine data. Verotta *et al.* (1991) generalised the model to accommodate possible irreversible blockade and noninstantaneous reversal kinetics and illustrated its ability to describe the time courses of edrophonium and neostigmine in humans. Thus PK/PD models can facilitate the cross-species comparisons of pharmacological activity.

An understanding of the time course of drug effect and how it is influenced by formulation, environmental, and patient factors is essential for the efficient design of clinical studies and the eventual use in therapy. In early clinical studies in humans, relevant dose-concentration-effect-time data are collected and analysed.

Subsequent intensive studies in appropriate patient groups using 'dose titration' or cross-over studies clarify individual dose-concentration-effect relationships over time, giving estimates of pharmacodynamic parameters. Pharmacodynamic modelling can provide important clues about the complexities of drug effect that only emerge over time such as sensitization, tolerance or the contributions of active metabolites. Early recognition of these characteristics can prevent irrational study designs that lead to inadequate demonstration of efficacy or excessive, unnecessary toxicity.

The total preclinical and clinical experience obtained during drug development should give rise to a reasonable understanding of the dose-concentration-effect-time relationship such that guidelines for individualisation of dosage regimens can be based on patient characteristics, disease intensity, and other risk factors (Sheiner, 1991). The power of PK/PD modelling to identify why important subpopulations of patients may respond differently has been illustrated for different drugs (Stanski and Maitre, 1990, Scott and Stanski , 1987).

## 1.2 Statistical Motivation and General Outline

The ultimate aim of this work is to describe a methodology to construct a model that both quantitatively and qualitatively associates drug administration to changes in complicated biological processes. The particular process we are going to focus on is the evolution of leukaemic cell populations. Therefore, the main goal will be to relate leukaemia treatment to the evolution of leukaemic cell populations. A better understanding of this relationship would improve current treatments by administering drugs more efficiently and effectively.

Any effort in trying to establish a direct and precise relationship between a drug dose and a number of leukaemic cells at certain time will most probably be condemned to failure due to the amount of intercorrelated biological and physiological processes, most of them unknown, involved. The attempt to encompass all these processes within a tractable statistical model without loosing its biological essence, is in itself of great interest.

The modelling approach taken in this work is primarily 'data-oriented' in the sense that the starting point of the task is to identify those relevant variables for which data can be collected. Blood samples are usually taken from patients during treatment. Standard techniques allow the measurement of blood drug concentration levels and the number of leukaemic cells in blood during treatment.

The structure of this work follows the natural steps required to build such a model based on surrogate information derived from blood samples: a pharmacokinetic model relating doses with blood drug concentration levels, a cell kinetic model explaining the evolution of leukaemic cell populations and a

pharmacodynamic model describing the effect that certain blood drug concentrations initiate in the kinetics of leukaemic cell populations. Necessary pharmacological concepts will be followed by the required statistical methodology and advances.

### Pharmacokinetics

All the physiological processes that determine the evolution of administered substances within the body should be captured by few manageable processes and adequate descriptive mathematical models built. This is discussed in chapter 2, where the main pharmacokinetic processes are explained and compartmental models introduced. These models aim to explain general trends of the corresponding processes.

Blood concentration levels and dosing histories of different patients are usually available. These data are usually very sparse both within and between individuals. In addition, repeated measures from the same individual are often correlated and conform to a pattern non-linear in a set of parameters of interest. These kind of scenarios have been successfully handled by setting appropriate 'non-linear hierarchical population models'. This approach will be taken in chapter 3 from a fully Bayesian approach that we advocate due to several beliefs and practical reasons that will be exposed. An introduction to Bayesian inference will be presented.

'Markov chain Monte Carlo' techniques facilitate enormously the inference for 'Bayesian non-linear hierarchical models'. The main algorithms and several technical and practical issues will be addressed and discussed.

The physiological and statistical theory developed in previous chapters will be applied to real data in chapter 4.

### Cell Kinetics

The second step required to build the aimed model is to study the evolution of leukaemic cell populations prior to treatment, that is, to study the cell kinetics of leukaemic cell populations. This will enable us to identify the drug-specific effect in the cell population.

Two approaches will be taken to model cell populations in chapter 5: the 'Continuous time Markov Process' approach and the 'Branching Process' approach. These two branches of stochastic modelling will be introduced and several new population-growth models built. The pros and cons of these models will be discussed. An easily tractable novel approximation of one of the main limit result of Branching Process theory will be presented.

### Pharmacodynamics

Lastly, when treatment starts the observed number of leukaemic cells will be modelled as a function of, apart from its inherent kinetics, blood drug concentration levels. The underlying processes are known to be drug and individual specific. General pharmacodynamic models are introduced in chapter 6. Observed data (number of leukaemic cells in blood samples) are usually very sparse and vary between and within individuals.

Consequently, in order to capture as much variability as possible, the population approach will be taken. Different new 'fully Bayesian non-linear population pharmacodynamic models' will be proposed for the special scenario we are concerned with, that is, changes in leukaemic cell populations due to treatment.

### Global Model

The resulting global model constitutes a novel 'fully Bayesian population pharmacokinetic-cell kinetic-pharmacodynamic hierarchical model' that, we believe, is able to capture the main underlying biological processes that dictaminate the effect of drugs.

This global model will be applied to real data in chapter 7. Unfortunately, as we will see, the resolution of the available data set is not the desired one, and several simplification the of the general model will be required. Therefore, the general model will not be fully tested in this work.

# Chapter 2

# Main Pharmacokinetic Processes

There are innumberable pharmacokinetic processes that will determine the fate of substances introduced in the organism. These can be grouped in three main processes: absorption, distribution and elimination.

The two main complementary approaches to analyse these processes quantitatively (*e.g.*, compartmental and non-compartmental approaches) are briefly introduced in the first section of the chapter. A qualitative and quantitative description of the absorption, distribution and elimination processes follow in sections 2.2, 2.3 and 2.4 respectively, using the language of the compartmental approach. Different mathematical models are built to describe the drug concentration time course within the body following a rapidly administered single dose, either intravenously or by the oral route.

More generalised dosing schemes are introduced in the pharmacokinetic models in the last two sections: infused dosing in section 2.5 and multiple dosing in section 2.6.

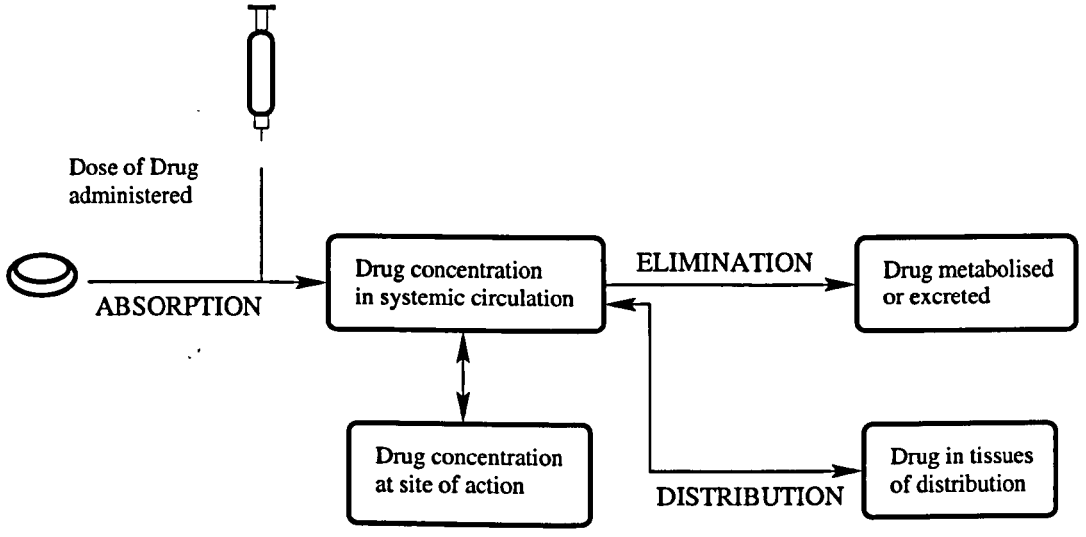
## 2.1 Introduction

There are two approaches to analysing pharmacokinetic data –usually blood drug concentration samples collected at different times. One is the 'compartmental analysis' where an appropriate mathematical model is built to describe the transfer of a drug from its absorption site to the blood and the various steps involved in the distribution of the drug throughout the body. This pharmacokinetic model has to be coherent with experimental data in order to use the achieved model parameters for prediction of blood plasma concentration after dosing. For compartmental pharmacokinetic theory and applications see, for instance, Gibaldi and Perrier (1982), Jusko (1992) and Rowland and Tozer (1995).

The second approach is the 'non-compartmental analysis'. Non-compartmental analysis does not require the assumption of a specific compartmental model for the drug. It rather focuses in the area under a plot of the plasma concentration versus time curve. The advantages of this approach are the disadvantages of the compartmental modelling approach and vice versa. Non-compartmental theory has been mainly developed in the last two decades due to the work of Benet and Galeazzi (1979), Watary and Benet (1989) and Nakashima and Benet (1989). For a discussion of the pros and cons of non-compartmental models versus compartmental models see, for example, Gillespie (1991), Kumperscak and Kozjek (1998) and Rescigno (2001).

These two approaches are not exclusive but complementary and should be used in conjunction for a better understanding of the kinetic processes that dictate the time course of the administered substance within the body. In what follows, we will concentrate on compartmental models, but considering them just as a part of the ultimate task. In fact, basic non-compartmental concepts will be introduced briefly within the compartmental framework.

Figure 2.1: Main Pharmacokinetic Processes. Schematic representation of the interrelationships between absorption, distribution and elimination of drugs that are injected or orally administered.




Therefore, to describe the fate of a drug in the body quantitatively, and even qualitatively, a mathematical model for the body is going to be built and assumed. To include all known kinetic processes in a pharmacokinetic model would be of no practical use because of mathematical difficulties. It is often more convenient to construct simple models that take into account only the main processes in a meaningful way. In the literature different compartmental models have been applied to study the pharmacokinetics of different drugs and dosing regimes. For recent applications, see for instance, Wu and Furlanut (1998), Lewitt and Lewitt (1998), McCarley and Bunge (2000) and Yukawa *et al* (2001).

The most important pharmacokinetic processes –e.g., Absorption, Distribution and Elimination– are shown schematically in figure 2.1. The administered substance needs first to be absorbed by the body before entering into the circulatory system. From the circulatory system drugs are both distributed to different parts of the body, including the site of action of interest, and

excreted irreversibly from the body. In what follows, each of these processes will be described in more detail.

## 2.2 Absorption

### 2.2.1 Introduction

As drugs are usually xenobiotics (i.e., substances that are chemically foreign to the body), they have to gain entrance into the body in order to produce an effect by the oral or the rectal route, inhalation, injecting it directly (i.e., intravenously) or indirectly (i.e., intramuscularly or subcutaneously) into the circulatory system , depending on the specific drug's characteristics and clinical targets. All these routes can be classified as 'enteral' (to do with the gastrointestinal tract, e.g., the oral and the rectal route) or 'parental' (other than intestine, e.g., intravenous, intramuscular or subcutaneous administration).

'Bioavailability' is defined as the fraction of unchanged or active drug reaching the systemic circulation following administration by any route, that is, the fraction which is absorbed by the organism. When drugs are administered parentally, the bioavailability is usually supposed to be equal to unity. Nevertheless, when orally administered, the dose is absorbed from the gastrointestinal tract into the circulation and, therefore, it can be extensively metabolised in the gut and in the liver before reaching the systemic circulation. This is called the 'first-pass effect' and it can be of great importance for some substances. This effect is drug dependent, and has to be studied in order to decide which drug to use and how to administer it for different specific targets.

Other factors which are of importance for the absorption extent of a drug are the chemical form (e.g., salt, ester), the dosage form (e.g., tablet, cap-

sule), and many irregular pharmacological or physiological processes which may complicate enormously the absorption process of a drug before reaching the circulatory system.

In addition to the fraction of the dose gaining entrance into the circulatory system, the velocity of the absorption process has to be studied as well. Slowly absorbed drugs may not achieve sufficiently high concentrations at the site of action to ensure the desired effect even if the entire dose is absorbed (e.g., unitary bioavailability factor). The bioavailability factor of a specific formulation of the drug will be one of the determinants of the therapeutic effect.

The 'Area Under the Concentration-time Curve', AUC, represents the amount of drug that is actually being absorbed. Hence, a comparison of these areas after equivalent oral and intravenous administration expressed as a percentage gives the bioavailability of a drug following oral administration.

Two typical drug concentration-time curves are shown in figure 2.2 for intravenous and oral dosing. Assuming that the administered doses are exactly the same in both cases, the bioavailability of this particular drug would be close to unity, as the areas under the curve do not seem to differ substantially.

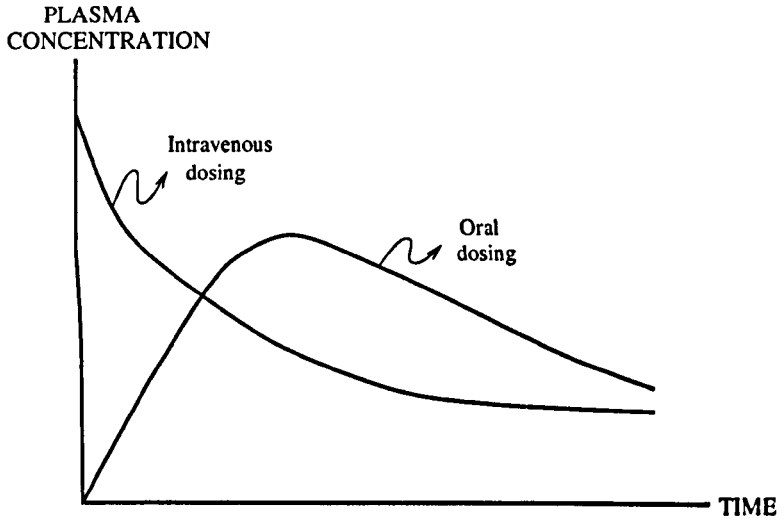
Note that concepts such as bioavailability and the AUC are purely physiological and that they do not rely on any model assumption.

In the next section a compartmental approach is presented to describe the absorption process where modelling assumptions are required and the concepts presented therein are, therefore, intrinsically model dependent.

### 2.2.2 Compartmental Approach

For many drugs, when orally administered, the absorption process from the gastrointestinal tract may be described by a 'first-order kinetic' process. With

Figure 2.2: Bioavailability of a drug. Comparison of concentration-time curves for intravenous and oral dosing.




first-order kinetics it is meant that the rate of change of drug concentration by any process is directly proportional to the drug concentration remaining to undertake that process. This concept will be explained with more detail in following sections.

With first-order absorption, the amount of drug in the gastrointestinal tract, $x_0$, may be described by the following linear differential equation:

$$\frac{dx_0}{dt} = -k_a x_0(t), \tag{2.1}$$

where $x_0(0)$ is given by the administered dose and $k_a > 0$ is the 'absorption rate constant'. This is the proportion of the remaining amount of drug in the gastrointestinal tract that enters into the circulatory system, which in a first-order absorption process is assumed to be constant.

This kinetic behaviour is linear, a simplification that in most applications has been done but that by no means should be generalised. Time dependency in kinetic variables has been observed for some drugs, and several non-linear models have been proposed and used in the literature. For general readings

on non-linear kinetic models see, for example, texts such as Lassen and Perl (1979), Gibaldi and Perrier (1982) and Levy (1982).

Equivalently, if we assume for the moment that the drug is not being eliminated from the body, the amount of drug in the circulatory system, say $x_1$, would then be given by the following linear differential equation:

$$\frac{dx_1}{dt} = k_a x_0(t), \tag{2.2}$$

where $x_0$ is the amount of drug in the gastrointestinal tract and $k_a$ is the absorption rate constant. The solution of equations (2.1) and (2.2) would determine the time course of the amount of drug in the gastrointestinal tract and in the circulatory system at any time.

In addition, in some applications the time course of concentration in blood may suggest a 'time-lag' between oral administration and the apparent onset of absorption. In this case equation (2.1) can be written as follows:

$$\frac{dx_0}{dt} = \begin{cases} 0 & \text{for } t < T_{\text{lag}} \\ -k_a x_0(t - T_{\text{lag}}) & \text{for } t \geq T_{\text{lag}} \end{cases}, \tag{2.3}$$

where $T_{\text{lag}}$ is the corresponding time-lag.

For further models and absorption routes see, for instance, Creasy and Jaffe (1991), Robinson (1991) and Andersson *et al* (1994).

## 2.3 Distribution

### 2.3.1 Introduction

Once a certain fraction of the administered drug has entered the circulatory system, blood samples can be taken and drug concentration in blood plasma

reported. A fraction of this drug will be bound to plasma proteins and will therefore be pharmacologically inactive. Therefore, all the processes involved in protein binding will be crucial for the amount that actually is free to be distributed along the tissues and, eventually, reaches the target site of action.

The volume of distribution, $V$, for a drug relates the amount of drug in the body, say $x$, to the concentration of drug, $C$, in blood plasma:

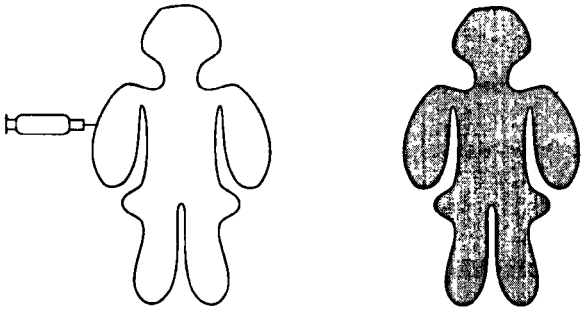
$$V = \frac{x}{C}. \tag{2.4}$$

The volume calculated using equation (2.4) is an 'apparent volume of distribution' because it does not refer to any identifiable physical site in the body. It is simply the volume necessary to contain the amount of drug homogeneously throughout the body at the concentration found in the blood plasma. Drugs with a very high apparent volume of distribution have much higher concentrations in extravascular tissues than in the vascular compartment and, therefore, they are not homogeneously distributed.

From the circulatory system, the drug moves to different anatomical tissues and organs following rather complex and sometimes correlated kinetics – distribution, metabolism, excretion. In the compartmental approach each of these sites, and even parts of them, are thought of as a single compartment where the substance enters at some point, stays in there for a certain period and leaves it reversibly or irreversibly, changed or unchanged, following certain kinetic behaviour. Within a compartment the kinetic mechanisms are supposed to be equal so that, the drug fraction that reaches a compartment is distributed uniformly along it, and it is eliminated or transferred to other compartments following the same kinetics all over it.

Every attempt for first identifying all the different compartments and second assigning specific kinetic patterns to each of them, seems therefore condemned

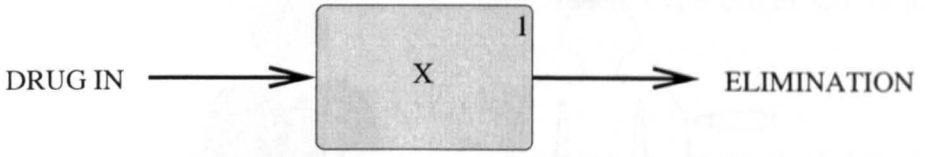
Figure 2.3: One-Compartment Model. The body is treated as a single compartment where the drug is intravenously administered. The drug is uniformly mixed in the body immediately after administration.




to failure. Nevertheless, quite often many of the processes involved in the evolution of a substance within the body might be fast or not significant for a given drug, being reasonable, and sometimes necessary, to model them as a single process.

In a multi-compartment model, drugs first enter into the central compartment, normally comprising the blood plasma, and from there are transferred reversibly into all other compartments, called peripheral compartments, following certain kinetic patterns. The actual identification of pharmacokinetic compartments with real anatomical tissues or organs is rather complex and, at times, impossible. When drugs are administered in a way that do not give direct entry into the central compartment (e.g., oral administration) the preliminary absorption phase needs to be introduced in the model as well. The absorption phase will take place in the corresponding absorption compartment.

Figure 2.4: One-Compartment Model. Substances enter the body which is understood as a single compartment where drugs enter and leave following a uniform kinetic pattern.




### 2.3.2 Single-Compartment Models

In a one-compartment model, the body is understood as a single compartment or liquid container where once a drug is introduced it spreads rapidly uniformly through all the organs and tissues. Figure 2.3 represents this first simple environment which is meaningful for certain drugs for which is assumed that all body compartments are in rapid equilibrium with the central compartment. This means that the drug is uniformly distributed throughout the body.

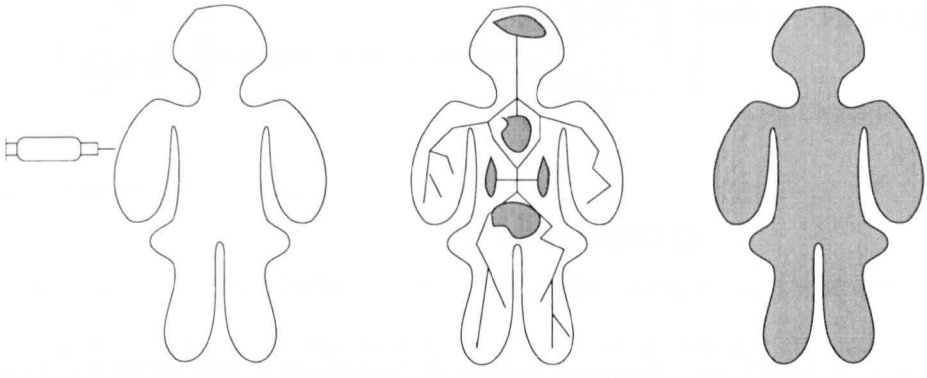
Figure 2.4 describes schematically the one-compartment model where the drug is injected directly into the central compartment and it is distributed instantaneously and uniformly through the only compartment that constitutes the whole body. Eventually the drug has to be eliminated from this only compartment.

Even though the one-compartment model might seem to be too simplistic, it has been applied successfully in some cases (Matucci *et al*, 1983, Edgren *et al*, 1984, Lee *et al*, 1989, Norman, 1992) and work in the methodology continues (Hoke and Ravis, 1992, Purves, 1993, Breant *et al*, 1996, Wu, 2000).

### 2.3.3 Multi-Compartment Models

Sometimes it is more convenient to conceptualise the body as two (or more) different compartments. The first compartment can be thought of as a rapid

Figure 2.5: Two-Compartment Model. The body is divided in two different compartments. Immediately after intravenous administration, the drug enters the circulatory system and organs and tissues with high blood-flow. More time is needed for the drug to mix in other organs and tissues that form the peripheral compartment.

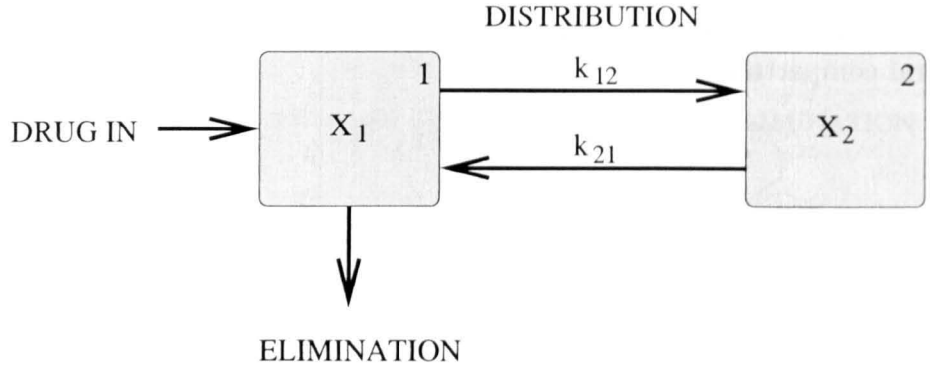



mixing central compartment, usually made up by blood and those organs or tissues that have high blood-flow. The second or peripheral compartment will equilibrate with the drug over a longer time period after administration. This is illustrated in figure 2.5.

When a two-compartment model is assumed, the central compartment has a volume referred to as $V_1$ or initial volume. The substance will then be distributed into the second or peripheral compartment, so that the initial volume will decrease as the volume in the second or peripheral compartment, $V_2$, increases. The sum of $V_1$ and $V_2$ is the total apparent volume of distribution, $V$. Therefore, high drug concentrations in plasma immediately after administration may be due to the time that consumes the deliberation of the drug into the peripheral compartment.

When designing a dose it is important to know whether the target site behaves as belonging to the central compartment or to the peripheral compartment, and to calculate it on the basis of the total volume of distribution, $V$. Plasma

Figure 2.6: Two-Compartment Model. There are two compartments with different kinetic patterns with constant flows between compartments.




concentrations which are obtained before distribution is complete will not reflect the peripheral compartment at equilibrium.

Figure 2.6 describes a two-compartment model where the drug is exclusively eliminated from the central compartment. Here, $k_{12}$ accounts for the fraction of drug that is distributed from the central compartment to the peripheral compartment, and $k_{21}$ for the fraction of drug that returns from the peripheral compartment to the central compartment. Usually these flows satisfy first order kinetics so that all the fractions are constant.

Two-compartment models are widely used in pharmacokinetics research. Recent applications and developments include Yashiro *et al* (1994), Smye and Will (1995), Chappell (1995), Artemov *et al* (1998), Tsoukias and George (1998), Yukawa *et al* (2001) and Catlin *et al* (2001).

In the same way, it is sometimes necessary to conceptualise three different compartments within the body when modelling plasma concentrations of certain drugs. Again, there will be a central compartment with rapid drug mixing, and two peripheral compartments where the mixing follows different kinetic patterns.

Three-compartment models are also being applied in pharmacokinetic studies of different drugs. See, for instance, Tingle and Park (1993), Sabot *et al* (1995) and Wu and Furlanut (1998) and references therein.

## 2.4 Elimination

### 2.4.1 Introduction

After, and during, a drug is distributed within the body it has to be eliminated. The fraction of drug that has not been absorbed is not considered as being eliminated, as it is considered to be 'outside' the body in the first place. There are some mechanisms whereby the molecule has its bioactivity terminated by biotransformation (metabolism) and is excreted from the body (through kidneys, liver, etc.). Otherwise, drug effects would last the lifetime of the host.

The two major sites of drug elimination are the kidneys and the liver. Clearance of unchanged drug in the urine represents 'renal clearance'. Within the liver, drug elimination occurs via biotransformation of parent drug to one or more metabolites, or excretion of unchanged drug into the bile, or both.

Drug clearance represents the volume of blood or plasma that is being completely cleared (because of elimination from the kidneys) of drug per unit of time,

$$C_L = \frac{R_E}{C}, \tag{2.5}$$

where $C$ is the plasma concentration and $R_E$ is the rate of elimination which accounts for the total amount cleared from the body per unit of time (not just from the plasma). The greater the value of plasma clearance, the greater will be the rate at which the drug will be removed from the body.

The total body clearance, $C_L$, comprises the renal clearance, the metabolic clearance and other possible clearance sites.

The rest of this section will be devoted to pharmacokinetic compartmental models following a single dose with no previous dosing.

### 2.4.2 The One-Compartment Model Following Intravenous Administration

Different types of drugs follow distinct kinetic patterns. When is said that the elimination process from the body of a drug follows a 'zero order kinetics' it is meant that the plasma concentration decreases at a fixed amount per time. A plot of concentration against time would produce a straight line.

Nevertheless, for many drugs the elimination process follows 'first order kinetics'. As said before, in a first-order kinetic process the rate of change in drug amount is proportional to the first power of the concentration of a drug. Therefore, the concentration of drug in the body will diminish logarithmically over time. This assumption implies that the rate of elimination, $R_E$,

$$R_E = -\frac{dx}{dt} = k_{el}x, \tag{2.6}$$

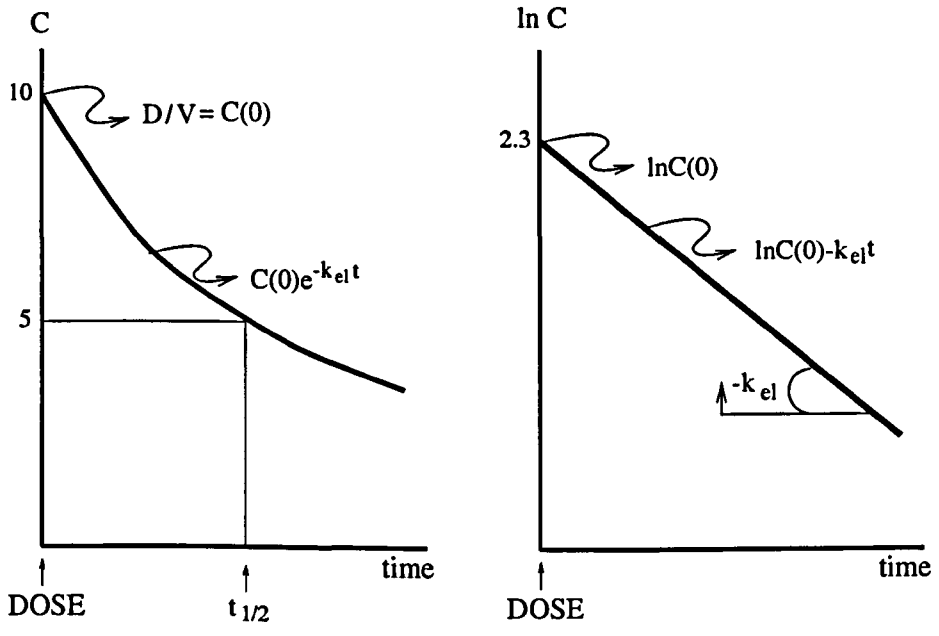
where $x$ is the total amount of drug in the body as illustrated in figure 2.7. The rate of elimination is proportional to the drug concentration and, therefore, the fraction of the total amount of drug that is removed per unit of time will remain constant, independently of the dose administered. This constant is called the 'Elimination Rate Constant', $k_{el}$, and satisfies:

$$k_{el} = \frac{R_E}{x} = \frac{R_E}{VC} = \frac{C_L C}{VC} = \frac{C_L}{V}. \tag{2.7}$$

We can now rewrite equation (2.5) and find the relationship between the elimination rate constant and the evolution of drug concentration:

$$k_{el} = -\frac{dx}{dt}/x = -\frac{dx}{dt}/(VC) = -\frac{dC}{dt}/C, \tag{2.8}$$

Figure 2.7: First order elimination in a one-compartment model following intravenous administration. Typical concentration-time and logarithmic concentration-time curves when a single dose is administered at time 0.




and consequently:

$$\frac{dC}{dt} = -k_{el}C(t). \tag{2.9}$$

This exponential behaviour of the plasma concentration is shown in figure 2.7. The elimination rate constant is the constant slope of the logarithm of the plasma concentration against time.

Integrating up equation (2.9), a single exponential equation follows:

$$C(t) = Ae^{-k_{el}t}, \tag{2.10}$$

where $A$ is the dose-dependent constant accounting for the concentration just immediately after administration. This constant can be determined by noting that, when no previous doses have been administered, immediately after intravenous administration the total amount of drug in the body is the dose (bioavailability equals to one when considering intravenous administration),

and therefore recalling equation (2.4),

$$V = \frac{D}{C(0)} \quad \Rightarrow \quad C(0) = \frac{D}{V} \quad \Rightarrow \quad C(t) = \frac{D}{V}e^{-k_{el}t}, \tag{2.11}$$

where $D$ is the administered dose.

For convenience, the elimination rate constant is often expressed in terms of the drug's 'elimination half-life', or '$\beta$ half-life'. The $\beta$ half-life of a drug is the amount of time required for the plasma drug concentration to decrease by one half starting from any time point. This concept is related to the elimination rate constant by the formula:

$$\frac{C(t)}{2} = C(t_{1/2}) = C(t)e^{k_{el}t_{1/2}} \Rightarrow \tag{2.12}$$

$$\ln 2 = k_{el}t_{1/2} \Rightarrow \tag{2.13}$$

$$t_{1/2} = \frac{\ln 2}{k_{el}} = \tau_{el}\ln 2, \tag{2.14}$$

where $\tau_{el}$ is the time constant for elimination that satisfies:

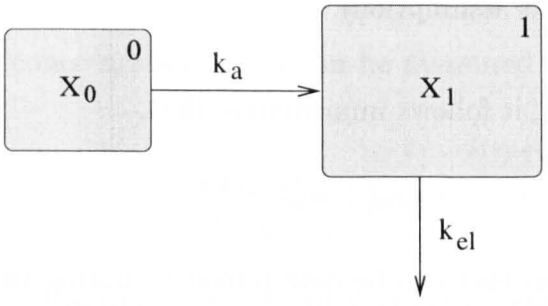
$$\tau_{el} = \frac{1}{k_{el}} = \frac{V}{C_L}.$$

Equation (2.11) can now be expressed as follows,

$$C(t) = \frac{D}{V}e^{-k_{el}t} = \frac{D}{V}e^{-\frac{t}{\tau_{el}}} = \frac{D}{V}2^{-\frac{t}{t_{1/2}}}. \tag{2.15}$$

This equation represents the relationship between the administered dose and the drug concentration. When the elimination half-time of a drug is known and there is a good understanding of the parameter $A$, it becomes possible to predict the drug concentration in the body at any time $t$ after dosing. It will take one half-time to eliminate 50% of all the drug administered, two half-times to eliminate 75%, three half-times to eliminate 87.5%, four half-times to eliminate 93.75% and so on.

Figure 2.8: One-Compartment Model after oral administration. The substance reaches the circulatory system at a constant rate $k_a$ and is eliminated from the body at rate $k_{el}$.




### 2.4.3 The One-Compartment Model Following Oral Administration

When a drug is orally administered it does not reach the circulatory system immediately after administration. As mentioned before, the substance has to be metabolised in the gastrointestinal tract and other tissues in order to be absorbed at some extent.

We will assume that drug absorption from the gastrointestinal tract may be described by a first-order kinetic process. That is to say that the amount of drug that is being absorbed by the gastrointestinal tract is proportional to the amount of drug remaining to undertake the absorption process (see equation 2.1). This constant proportion is called the 'absorption rate constant' $k_a$, already introduced.

The differential equations that describe first-order absorption and elimination represented in figure 2.8 are the following:

$$\frac{dx_1}{dt} = -k_{el}x_1(t) + kax_0(t), \tag{2.16}$$

$$\frac{dx_0}{dt} = -k_a x_0(t), \tag{2.17}$$

where $x_0$ is the total amount of drug that enters the body (e.g., the absorption compartment) and $x_1$ accounts for the amount of active drug that enters the body (e.g., the circulatory system and all the body organ and tissues within the monocompartment assumption).

From equation (2.17), it follows immediately that,

$$x_0(t) = A_1 e^{-k_a t}, \tag{2.18}$$

where $A_1$ is a constant that can be determined by noting that right after oral administration the total amount in the absorption compartment is the fraction of the dose that reaches the absorption phase actively. That is,

$$x_0(0) = A_1 = FD, \tag{2.19}$$

where $F$ denotes the bioavailability factor and $D$ is the administered dose. Sometimes, as it will be shown, is necessary to consider a short delay between the intake of a medicine and the onset of the absorption process.

From equation (2.19), it follows immediately that,

$$x_0(t) = FDe^{-k_a t}. \tag{2.20}$$

It is straightforward now to solve equation (2.16). Substituting $x_0(t)$ in equation (2.16) we get that,

$$x_1(t) = FD\frac{k_a}{k_{el} - k_a}e^{-k_a t} + A_2 e^{-k_{el} t}, \tag{2.21}$$

where $A_2$ is a constant to be determined. Assuming that at the time of administration there is no drug in the body (that is to say that no drug has been administered previously), it follows that,

$$x_1(0) = FD\frac{k_a}{k_{el} - k_a} + A_2 = 0, \tag{2.22}$$

and therefore,

$$A_2 = FD\frac{k_a}{k_a - k_{el}}. \tag{2.23}$$

Hence, the solution is given by:

$$x_1(t) = FD\frac{k_a}{k_a - k_{el}}(e^{-k_{el}t} - e^{-k_a t}). \tag{2.24}$$

The real amount of active medicine in the body, $x_1$, is unobservable. Nevertheless, drug blood concentration levels can be measured and, from equation (2.4), it follows that,

$$C(t) = \frac{x_1(t)}{V} = \frac{FD}{V}\frac{k_a}{k_a - k_{el}}(e^{-k_{el}t} - e^{-k_a t}). \tag{2.25}$$

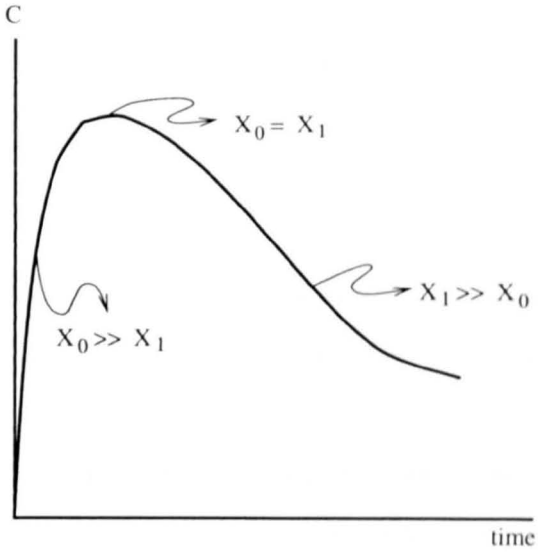
In figure 2.9 a theoretical drug concentration curve is represented. Generally $k_a > k_{el}$, and hence the second exponential of equation (2.25) is usually faster than the first one. Therefore, the concentration-time curve starts increasing exponentially during the absorption phase, reaches a peak value and finally decreases exponentially once most of the drug has been absorbed and is being eliminated from the body. This shape is characteristic of first-order absorption and elimination kinetics.

### 2.4.4 The Two-Compartment Model Following Intravenous Administration

As mentioned in previous sections, sometimes it is too naive to model the body as a single, rapidly mixing compartment. Most drugs follow very complicated and 'obscure' kinetics within the body. The main difficulty when dealing with complicated multi-compartment pharmacokinetic models is that it is very hazardous to draw conclusions on the drug behaviour within several compartments from measurements only in plasma, and sometimes urine, compartments.

Generally it is clinically expensive, and sometimes even impossible, to obtain the profile of drug concentration in other body fluids, tissues or organs. However, the dose-concentration relationship in different body compartments is, for many drugs, much more homogeneous than one would expect at first glance.

Figure 2.9: One-Compartment Model after oral administration. After administration, the substance is absorbed, $x_0$, and therefore plasma concentration, $x_1$, increases until drug elimination exceeds drug absorption.

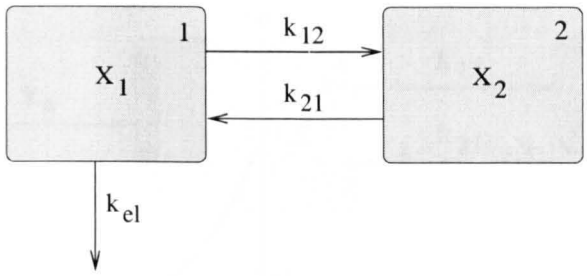



In a two-compartment model, two different kinetic patterns are modelled, one for the central compartment and one for the peripheral compartment. This is represented schematically in figure 2.10. Assuming first order kinetics within and between compartments, the dynamics governing the total amount of drug in the central compartment following intravenous administration can be described by the following equation:

$$\frac{dx_1}{dt} = -(k_{12} + k_{el})x_1(t) + k_{21}x_2(t), \tag{2.26}$$

where $x_1$ is the total amount of drug in the central compartment and $x_2$ is the amount of drug in the peripheral compartment. It is assumed that drug elimination (either by biotransformation or by excretion) occurs exclusively from the central compartment at a constant rate $k_{el}$. This assumption is usually made, as in most cases the vast mayority of the elimination process occurs from the central compartment. The substance flows from the central compartment to the peripheral one at a constant rate $k_{12}$ and it returns to the central compartment at a constant rate $k_{21}$.

Figure 2.10: Two-Compartment Model. Substances migrate at a constant rates $k_{12}$ and $k_{21}$ between compartments 1, or central, and 2, or peripheral. Elimination occurs exclusively from compartment 1 at a constant rate $k_{el}$.




Therefore, the amount of drug at the second or peripheral compartment satisfies the following kinetics:

$$\frac{dx_2}{dt} = k_{12}x_1(t) - k_{21}x_2(t). \tag{2.27}$$

Integrating up these equations we end up with the amount of drug at the central compartment as a exponential function of time:

$$x_1(t) = A_1^* e^{-a_1 t} + A_2^* e^{-a_2 t}, \tag{2.28}$$

where the constants $A_1^*$, $A_2^*$, $a_1$, $a_2$, are determined by the initial conditions.

As $C(t) = x_1(t)/V_1$, where $V_1$ is the volume of distribution at the first compartment, it follows that:

$$C(t) = A_1 e^{-a_1 t} + A_2 e^{-a_2 t}, \tag{2.29}$$

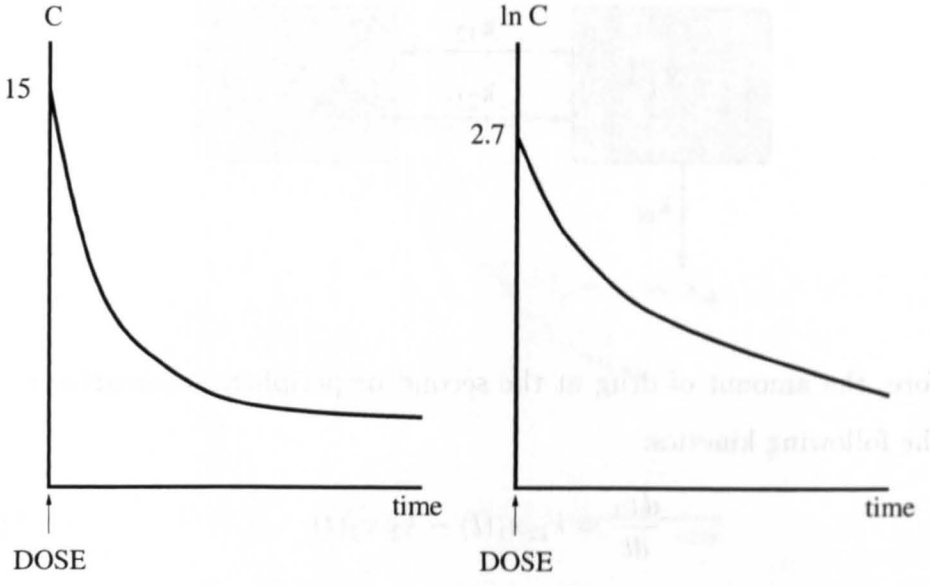
where $A_1 = A_1^*/V_1$ and $A_2 = A_2^*/V_1$. It is straightforward to see that the constants $A_1, A_2, a_1, a_2$, in terms of the original parameters are given by:

$$A_1 = \frac{D(a_1 - k_{21})}{V_1(a_1 - a_2)}, \tag{2.30}$$

$$A_2 = \frac{D(k_{21} - a_2)}{V_1(a_1 - a_2)}, \tag{2.31}$$

where $a_1$ and $a_2$ $(a_1 > a_2)$ are the roots of the quadratic equation: $r^2 + (k_{12} + k_{21} + k_{el})r + k_{21}k_{el} = 0$.

Figure 2.11: Two-Compartment Model after intravenous administration. Typical concentration-time and logarithmic concentration-time curves when a single dose is administered at time 0 in a two-compartment model.

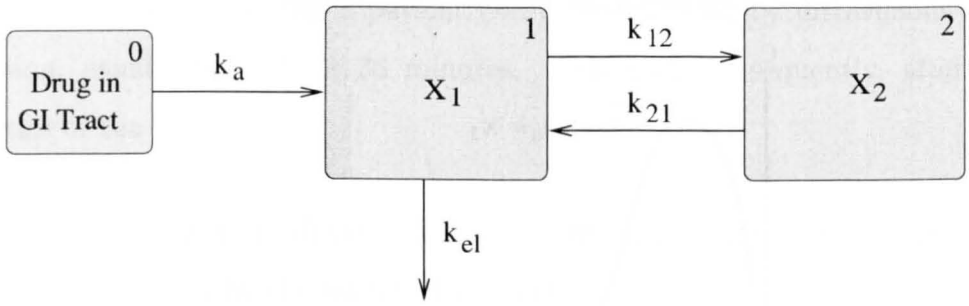



A typical exponential behaviour of the concentration-time curve derived from equation (2.29) is presented in figure 2.11. In the right plot the logarithmic concentration-time curve is presented. Comparing figures 2.11 and 2.7, we can see that in a two-compartment model the concentration in blood, or first compartment, decreases more rapidly immediately after administration due to drug distribution into the peripheral compartment. Once the elimination becomes the dominant process, the logarithmic concentration-time curve will be approximately linear as in the single compartment case shown in figure 2.7.

### 2.4.5 The Two-Compartment Model Following Oral Administration

It has been explained in section 2.2 that when drugs are administered orally, the absorption process needs to be modelled. Assuming first-order absorption,

Figure 2.12: Two-Compartments Model after oral administration. The substance is absorbed in the gastrointestinal tract (GI tract) and enters the central compartment, $X_0$, at a constant rate $k_a$. Distribution and elimination occurs at rates $k_{12}$, $k_{21}$ and $k_{el}$.




the dynamics of the two-compartment model following oral administration are given by the following set of differential equations:

$$\frac{dx_1(t)}{dt} = -(k_{el} + k_{12})x_1(t) + k_a x_0(t) + k_{21}x_2(t), \tag{2.32}$$

$$\frac{dx_0(t)}{dt} = -k_a x_0(t), \tag{2.33}$$

$$\frac{dx_2(t)}{dt} = k_{12}x_1(t) - k_{21}x_2(t), \tag{2.34}$$

where the notation has already been introduced. This kinetic behaviour is schematically shown in figure 2.12.

Solving equations (2.34)-(2.36), it follows that,

$$x_1(t) = A_1^* e^{-a_1 t} + A_2^* e^{-a_2 t} + A_3^* e^{-k_a t}, \tag{2.35}$$

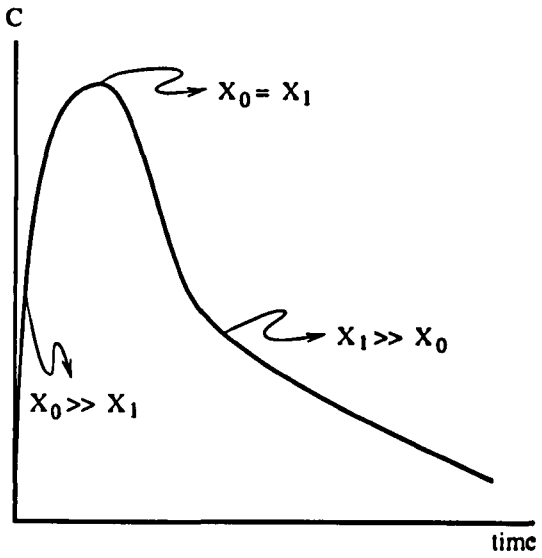
where the constant $A_1^*, A_2^*, A_3^*, a_1$ and $a_2$ in terms of the original parameters are given by:

$$A_1^* = \frac{Dk_a(k_{21} - a_1)}{(a_1 - a_2)(a_1 - k_a)}, \tag{2.36}$$

$$A_2^* = \frac{-Dk_a(k_{21} - a_2)}{(a_1 - a_2)(a_2 - k_1)}, \tag{2.37}$$

$$A_3^* = \frac{Dk_a(k_{21} - k_a)}{(a_2 - k_a)(a_1 - k_a)}, \tag{2.38}$$

Figure 2.13: Two-Compartment Model after oral administration. After administration, the substance is absorbed and, therefore, plasma concentration increases until drug elimination exceeds drug absorption. The substance is then distributed between the central and the peripheral compartments. Elimination occurs exclusively from the central compartment.




where $a_1$ and $a_2$ $(a_1 > a_2)$ are the roots of the quadratic equation: $r^2 + (k_{12} + k_{21} + k_{el})r + k_{21}k_{el} = 0$.

As the concentration is defined as $C(t) = x_1(t)/V_1$, it follows that,

$$C_1(t) = A_1 e^{-a_1 t} + A_2 e^{-a_2 t} + A_3 e^{-k_a t}, \tag{2.39}$$

where, $A_1 = A_1^*/V_1$, $A_2^* = A_2/V_1$, and $A_3^* = A_3/V_1$.

Figure 2.13 shows a typical blood drug concentration - time curve following the exponential behaviour derived from equation (2.45). The concentration in blood increases exponentially during the absorption phase. Comparing figures 2.13 and 2.9, notice that after the drug is absorbed, blood concentration decreases initially more rapidly in this two-compartment model than in the one-compartment model. This is due to the distribution process between the central and the peripheral compartments.

## 2.5 Infusion

In previous sections, drugs have been administered either orally, in tablet form, or by rapid intravenous injection, known as a 'bolus' dose. Nevertheless, commonly in a hospital setting a patient will receive a drug by intravenous slow infusion, usually over 15 or 30 minutes. This will, consequently, alter the kinetics of the drug.

Let $T^I$ be the length of infusion. The 'constant infusion rate', $R$, will be equal to the dose divided by the length of infusion.

In the one-compartment model it is easy to see that (see, for example, Gibaldi and Perrier, 1982) the concentration of drug in the blood following a single intravenously infused dose is given by:

$$C(t) = \frac{D}{T^I}\frac{1}{Vk_{el}}(e^{-k_{el}t'} - e^{-k_{el}t}), \tag{2.40}$$

where $T^I$ is the length of infusion, $t' = t - T^I$ for $t > T^I$ and $t' = 0$ otherwise, $D$ is the dose, $V$ represents the volume of distribution and $k_{el}$ the elimination rate constant.

When the kinetics of the drug are not uniform throughout the body and a two-compartment model is required, the corresponding concentration following a single intravenously infused dose will be given by:

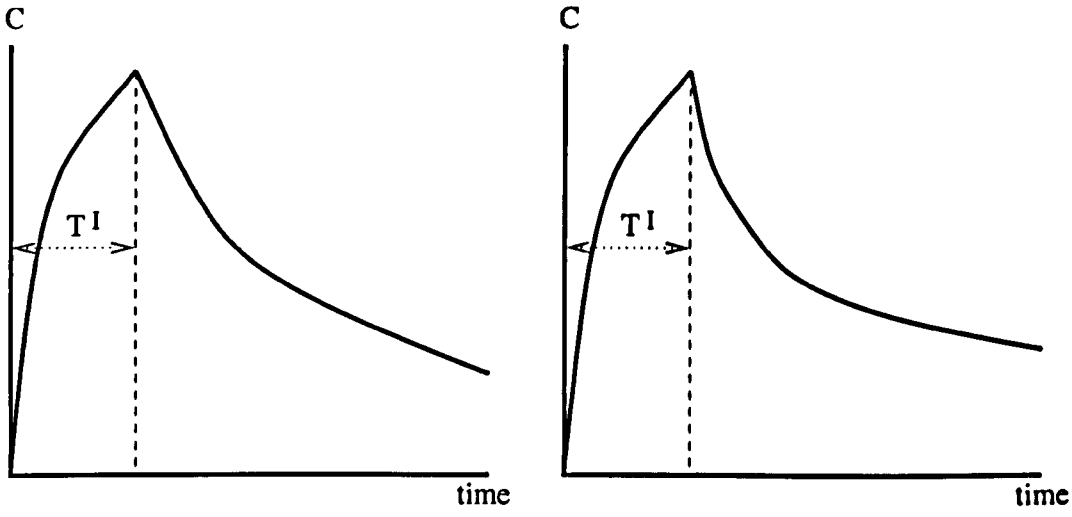
$$C(t) = A(e^{-a_1 t} - e^{-a_1 t'}) + B(e^{-b_1 t} - e^{-b_1 t'}), \tag{2.41}$$

where $T^I$ is the length of infusion, $t' = t - T^I$ for $t > T^I$ and $t' = 0$ otherwise. The constants $A$ and $B$ when no previous doses have been administered, are given by:

$$A = \frac{D}{T^I V_1}\frac{(k_{21} - a_1)}{(a_1 - a_2)a_1}, \tag{2.42}$$

$$B = \frac{-D}{T^I V_1}\frac{(k_{21} - a_2)}{(a_1 - a_2)a_2}, \tag{2.43}$$

Figure 2.14: Constant-rate infusion. A drug is infused at a constant-rate during the length of infusion, $T^I$, followed by a one-compartment kinetic behaviour (left plot) and by a two-compartment kinetic behaviour(right plot).



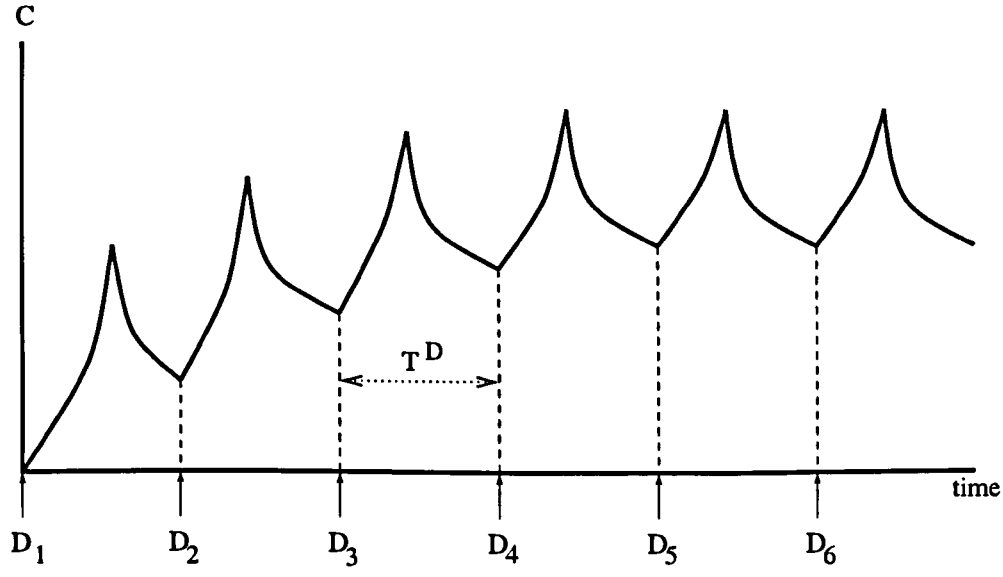

where, $a_1$ and $a_2$ $(a_1 > a_2)$ are the roots of the quadratic equation: $r^2 + (k_{12} + k_{21} + k_{el})r + k_{21}k_{el} = 0$. The rest of the notation has already been introduced in previous sections.

Typical concentration-time plots for one- and two-compartment models following a single infused dose are shown in figure 2.14. Drug is being eliminated from the start of the infusion, but as the infusion rate exceeds the elimination rate, concentration levels increase during the infusion period. When the infusion is stopped, the concentration will decrease according to a one-compartment kinetic (left plot) or to a two-compartment kinetic (right plot) behaviour.

## 2.6 Multiple Dosing

Different pharmacokinetic models that can explain the evolution of a drug in the circulatory system following a single dose administered either by an intra-

Figure 2.15: Multiple dosing. Six doses of a drug are infused intravenously at constant intervals $T^D$. The dosing regime has been designed so that after the fifth dose, a steady-state drug concentration is achieved. By 'steady-state' is meant that the concentration at any time during any dosing interval is identical to the concentration at the same time during any other interval.




venous bolus or infusion or by the oral route have been described in previous sections. Some drugs may be effective as a single dose, for example analgesics. Nonetheless, more frequently, drugs are administered on a continuous basis.

Usually, multiple doses are given at fixed intervals with sufficient frequency that significant levels of drug persist in the body at dosing times. The peak concentration level following the second and succeeding doses is frequently higher than the peak level after the first dose. This is not due to higher doses but due to the fact that concentration levels from previous doses accumulate in the body.

In fact, generally the dosing regime is designed in such a way that drug accumulation proceeds at a decreasing rate with increasing number of doses until a 'steady-state' concentration level is achieved. At steady state, drug concentration in blood at any time during any dosing interval should be identical to

the concentration at the same time during any other dosing interval.

Figure 2.15 shows a drug concentration time course due to a treatment consisting on multiple doses of a drug at fixed time intervals. The doses are in this case calculated in order to achieve steady-state drug concentrations after the fifth dose is administered. Therefore, denoting by $t_{D_n}$ the time of the $n^{\text{th}}$ dose, it follows that:

$$C(t_{D_5} + l) = C(t_{D_n} + l), \tag{2.44}$$

for all $l \geq 0$ and $n > 5$.

The concept of steady-state drug concentration levels constitutes one of the bases of a rational treatment. We are not going to talk any further about it because it will not be required in subsequent chapters. For further reading, see for instance, the book by Gibaldi and Perrier (1982).

## 2.7 Discussion

The basic pharmacokinetic processes have been explained to sufficient extent to derive the most common compartmental models understanding the underlying physiological mechanisms.

As has been shown, scientists comprise complicated physiological processes in tractable mathematical models. The aim is to obtain rigorous explanatory, descriptive and/or predictive information about the general trends of the main pharmacokinetic processes for particular drugs administered in different manners.

# Chapter 3

# Population Pharmacokinetics

In a standard pharmacokinetic study individual dose histories, individual covariate information and blood drug concentration measures with associated sampling times are usually reported. The aim is to fit the most appropriate, both physiologically and statistically, pharmacokinetic model. Pharmacokinetic data are usually very sparse both between and within individuals and two approaches can be taken to estimate the pharmacokinetic parameters: individual and population approach. The second is taken in this work due to the firm belief that there exists a high correlation structure between individuals' parameters. This constitutes valuable information that should by no means be omitted when estimating and/or predicting individual parameters.

All the pharmacokinetic models described in the previous chapter are nonlinear in the parameters and the dimensionality of the parameter space is usually large. For these cases, classical inference tools rely on asymptotic arguments that can be hazardous to examine for particular data sets. A fully Bayesian approach is taken in this work instead. Due primarily to spectacular developments in 'Markov chain Monte Carlo' techniques in the last decade, the Bayesian approach offers a flexible framework to build and estimate population pharmacokinetic models that can be easily used for prediction and, therefore,

dose recommendation. In addition, when relevant prior information from other studies or from expert advise is available and correctly used, inference can be considerably improved.

Population pharmacokinetic models and Bayesian statistics and Markov chain Monte Carlo techniques will be reviewed in this chapter. A fuller treatment and results of the former can be found, for instance, in Wakefield and Racine-Poon (1995). The review of Bayesian statistics and Markov chain Monte Carlo techniques is not at all complete and can be found in general references such as Bernardo and Smith (1994) and Gilks *et al* (1996).

An introduction to population pharmacokinetic models in section 3.1 is followed by a general three-stage hierarchical model that accommodates all the models that will be applied in succeeding chapters. In section 3.3 a 'directed acyclic graph' is constructed for the general population pharmacokinetic model in order to specify the full joint distribution of all model quantities in an intuitive manner.

The rest of the chapter is devoted to a review of Bayesian inference. Special attention is given to Markov chain Monte Carlo tools that have contributed enormously to the applicability of Bayesian inference. Some of these tools will be used in following chapters.

## 3.1 Introduction

Population models have been widely used in biometrical curve analysis (Berkey, 1982), in educational research (Novick *et al.*, 1972), econometrics (Swamy, 1970) and many other fields. The problem of implementing such models attracts methodological attention in the statistical literature both from a non-Bayesian perspective (Lindstrom and Bates, 1990) and from a Bayesian per-

spective (Gelfand and Smith, 1990). We adopt the Bayesian methodology due to reasons that will be explained in following sections and chapters.

Population analyses attempt to explain and control the variability observed both within and across individuals. Kinetic data are usually very sparse because of the fact that in clinical trials various doses and/or dosage regimens and/or application routes (e.g. oral, intravenous, intramuscular) are generally used in different patients and in the same patient at different stages of the therapy. Also several responses might be measured (e.g., drug plasma concentration, arterial blood pressure), and diverse administration schedules (single dose and chronic dosing) might be considered.

Population sparse data arise specially in the later phases of drug development when the drug is administered to the patient population, usually a highly heterogeneous group. In early phases full profiles are available on a small number of well controlled individuals who are typically from a well-defined population.

The basic rationale behind population models is that the set of underlying structural pharmacokinetic parameters is qualitatively the same for all individuals and that what vary from individual to individual are the quantitative values.

These special characteristics of typical population data suggest that a population mathematical model should encompass a collection of models for heterogeneous individual observations.

The population approach has gained prevalence over the individual pharmacokinetic approach during the last two decades both from the Bayesian and from the non-Bayesian approach, and has proved enormous potential to deal with pharmacokinetic sparse data. The literature in this topic is very vast both

in the methodology and in particular applications (see, for instance, Steimer *et al*, 1994, Wakefield and Racine-Poon, 1995, Vozeh *et al*, 1996 and Racine-Poon and Wakefield, 1996, and references therein).

## 3.2 Hierarchical Structure

### 3.2.1 Introduction

Population Pharmacokinetic models fit naturally into a hierarchical structure (Wakefield and Racine-Poon, 1995). At the first stage of the hierarchy we model the data (e.g., blood drug concentration levels) from a particular individual over time as a function of unknown parameters.

At the second stage, the pharmacokinetic parameters from a particular individual over time as a function of unknown parameters, defining each of the individual's curves, are assigned some distributional form after accounting for covariate information, if any is going to be considered. A Bayesian model then requires a third stage at which prior distributions are specified for the hyperparameters defining the second stage distributional form.

Such a model acknowledges and estimates both the within-individual concentrations variability (at the first stage) and the variability observed between different individual's pharmacokinetic parameters (at the third stage). By doing this, it is possible to determine appropriate subject-specific dosage regimens that will lead to attainment of desirable concentration and response levels.

### 3.2.2 The General Population Pharmacokinetic Model

Each stage of the hierarchy is now described in more detail. We present a general model that accommodates most of the drug-specific models (Wakefield and Racine-Poon, 1995). The population pharmacokinetic models used in following chapters will all be particular cases of this general model.

The observed drug concentration of an individual within the framework of a general pharmacokinetic population non-linear mixed-effects model can be described as:

$$c_{ij} = f_{ij}(\theta_{ij}, t_{ij}, H_{ij}) + \epsilon_{ij} \tag{3.1}$$

where $c_{ij}$ for $j = 1, ..., n_i$ are the observed data (e.g., $n_i$ plasma concentration levels measured at time points $t_{ij}$) of the $i$th subject. An appropriate model of this type is defined for all $i = 1, ..., K$, where $K$ is the number of subjects in the sample. The function $f_{ij}$ is a specified pharmacokinetic function that predicts the concentration in the $i$th subject at time $j$, $\theta_{ij}$ is the vector of unknown subject-specific pharmacokinetic parameters (*e.g.*, volume of distribution and elimination rate constant), $H_{ij}$ describes the dose history up to time $t_{ij}$ (i.e. amounts and times of administration) and $\epsilon_{ij}$ accounts for the error between the 'true' (unknown) concentration and the corresponding measurement. Usually these errors are considered to arise from a distribution with zero mean and precision parameter $\tau_{ij}$. Generally, as discussed in the previous chapter, this Pharmacokinetic model will take the form of a sum of exponentials.

As mentioned before, various doses and/or dosage regimens can be used in different patients and in the same patient at different stages of the therapy. Accordingly, the functions $f_{ij}$ will sometimes differ within and across individuals.

To complete the first stage of the hierarchy we need to assign distributional forms to the errors, $\epsilon_{ij}$.

At the second stage of the hierarchy, a model relating the parameters of the different individuals is required. Let $x_{ij}$ represent the value of relevant pharmacokinetic covariates of the individual $i$ at times $t_{ij}$. The population approach assumes that after conditioning on covariates (in a parametric or non-parametric fashion), the Pharmacokinetic parameters are drawn from some assigned distribution.

A model commonly assumed at the second stage is given by

$$\theta_{ij} = g(\mu, x_{ij}) + \delta_{ij}, \tag{3.2}$$

where $\mu$ represents population pharmacokinetic parameters and $g(.)$ the value of these parameters predicted by the specific covariates of the individual, $x_{ij}$. Therefore, we have that

$$E[\theta_{ij}|x_{ij}] = g(\mu, x_{ij}). \tag{3.3}$$

The distribution of the 'random-effects' parameters $\delta_{ij}$ is defined by the population pharmacokinetic scale matrix $\Omega$.

To complete the specification of a full probability model, we require prior distributions for variables without parents or *founders* (e.g. $\tau_{ij}$, $\mu$, $\Omega$). We would often like these priors to be not too influential in the final conclusions, although if there is external information it may be very useful to be able to include it. In hierarchical models, it is particularly important to avoid casual use of standard improper priors since these may result in improper posterior distributions (see, for instance, DuMouchel and Waternaux, 1992).

The appropriate specification of 'non-informative' priors is an old problem in Bayesian statistics, and is particularly important when techniques are to be used in a scientific context in which an 'objective' inference is required. This will be further discussed in following sections.

In the present work, we will require that a full probability model is defined, and this forces all priors to be 'non-informative' but proper. It is, however, vital to distinguish parameters of primary interest from those which specify secondary structure for the model. The former will generally be location parameters, such as regression coefficients, and in many cases a normal prior with an extremely small precision (large variance) is adequate.

More care is required in setting the parameters that constitute secondary aspects of a model. The literature has recommended to think carefully about reasonable values for such parameters in advance and so specifying fairly informative priors wherever possible – the inclusion of such external information is unlikely to bias the scientific conclusions of a study, although some sensitivity analysis should be performed.

## 3.3 DAG representation

Figure 3.1 shows a *directed acyclic graph* (DAG) that corresponds to the model defined by equations (3.1)-(3.3) (*directed* because each link between nodes is an arrow; *acyclic* because by following the direction of the arrows, it is impossible to return to a node after leaving it). Each quantity in the model appears as a node in the graph, and directed links correspond to direct dependencies as specified above: solid arrows are probabilistic dependencies, while dashed arrows show functional (deterministic) relationships. The latter are included to simplify the graph but are collapsed over when identifying probabilistic relationships. Repetitive structures are shown as stacked 'plates'. Rectangles denote either quantities assumed to be fixed or observed data, and circles represent all unknown quantities.

It is important to understand that the graph represents properties of the full model before any data are observed and forms a convenient basis for the spec-

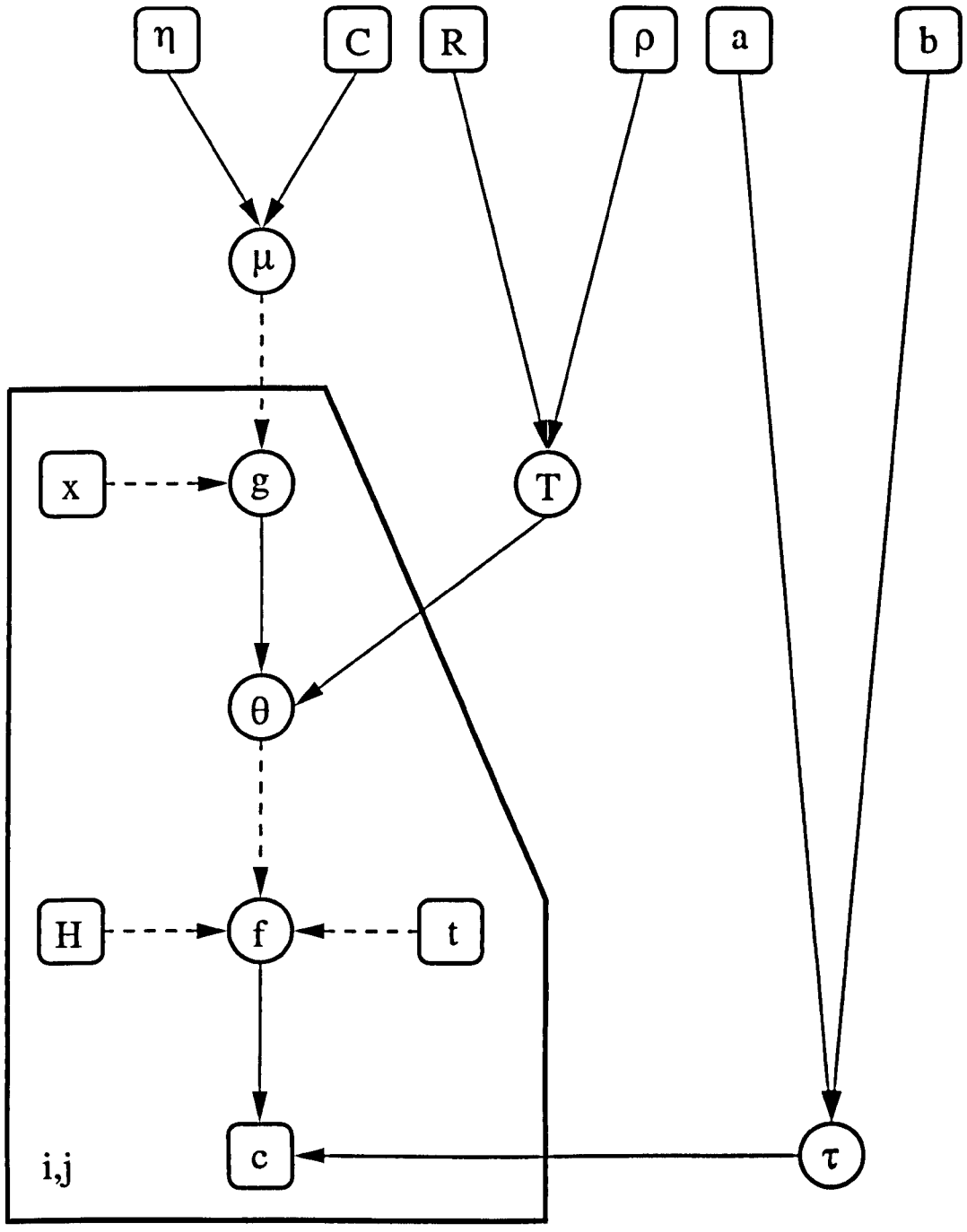



Figure 3.1: DAG representation for the General Pharmacokinetic Model. Solid and dashed arrows represent probabilistic and deterministic relationships respectively. Circles denote unknown random quantities whereas fixed or observed quantities are represented by rectangles. Stacked 'plates' represent repetitive structures.

ification of the full joint distribution of all model quantities; see for example Lauritzen *et al.* (1990) or Whittaker (1990) for discussion of how to read off independence properties from DAGs.

It can be shown (Lauritzen, 1996) that a DAG model is equivalent to assuming that the joint distribution of all the random quantities is fully specified in terms of the conditional distribution of each node given its parents:

$$P(V) = \prod_{v \in V} P(v|\text{parents}[v]) \tag{3.4}$$

where $P$ denotes a probability distribution, $v$ is a node in the graph, $V$ is the set of all nodes and a 'parent' of $v$ is any node with an arrow emanating from it pointing to $v$ – in identifying parents, deterministic links are collapsed. This factorisation not only allows extremely complex models to be built up from local components, but also provides an efficient basis for the implementation of some forms of MCMC methods that will be introduced in the following section.

## 3.4 Estimation

### 3.4.1 Bayesian Statistics

Until quite recently Bayesian statistics was undertaken largely by those primarily concerned with contributing to the theory, and the proportion of applied work that was formally Bayesian was rather small. There are several reasons for this. First, Bayesian statistics demands a prior distribution as well as a likelihood function, where non-Bayesian methods do not. Second, a subjective prior distribution has to be defended. Third, in most problems posterior distributions cannot be obtained anyway because the required integrals cannot be evaluated.

When no prior information is available, the idea of a 'non-informative' prior distribution, representing 'ignorance' and 'letting the data speak for themselves' in order to provide objective inferences sounds very seductive. A lot of work has been done to identify 'appropriate' procedures for formulating prior representations of ignorance, but no general procedure has yet been universally proposed. Bernardo and Smith (1994), provide an overview of some of the main directions followed in this controversial topic.

In many cases, however, prior information does exist. From a theoretical perspective discussions of the prior component inevitably focus on stylised forms, such as conjugate or reference specifications, which are amenable to a mathematical treatment. However, there is a danger of losing sight of the fact that, in real applications, prior specifications should be encapsulations of actual beliefs rather than mathematically 'nice' forms. The key task is how to process the expert knowledge and opinions in the form of probability distribution. There is a very vast literature in *Prior Elicitation.* See, for example, Merkhofer (1987), Garthwaite and Dickey (1992) or West and Crosse (1992).

Bayesian methods are operational only to the extent that posterior distributions can actually be computed. There are three ways in which this can be done. If the posterior distribution and the function of interest are sufficiently simple, the posterior distribution may be obtained analytically. Most results in this category may be found in Bernardo and Smith (1994). If the required integration takes place in few dimensions, then classical deterministic methods of numerical analysis, principally quadrature, are often practical. A standard reference for these methods is in Davis and Rabinowitz (1984). In the remaining cases, which constitute the preponderance of applied work, the approach of choice is posterior simulation using methods such as 'Rejection Sampling', 'Importance Sampling' and 'Markov chain Monte Carlo'.

The idea behind rejection sampling is to generate a random vector from a distribution that is similar, in an appropriate sense, to the posterior distribution, and then to accept that drawing with a probability that depends on the drawn value of the vector or otherwise repeat. If this acceptance probability function is chosen correctly the accepted values will have the desired distribution.

In importance sampling, however, rather than accept only a fraction of the draws from the source density, it retains all of them, and consistently approximate the posterior moments by appropriately weighting the draws.

The ability to generate draws from non-standard posterior distributions has improved enormously in the past ten years, due to the development of Markov chain Monte Carlo (MCMC) methods and the dramatic decrease in the cost of computing.

The development of MCMC methods in the last decade enables the simulation of samples from high-dimensional joint density functions and, as a result, the Bayesian approach offers a flexible and unified way of combining prior and sample information to obtain a complete posterior distribution upon which to base the inferences [1]. MCMC methods will be introduced in subsequent sections.

### 3.4.2 Bayesian Inference

This section provides a brief overview of Bayesian inference in order to introduce an explicit context of concepts and notation for following sections. The concepts will be very standard and can be found in many reference books, including Berger (1985) and Bernardo and Smith (1994).

[1] MCMC methods are also valuable in frequentist inference to explore the likelihood function.

Inference takes place in the context of one or more models. A model describes the behaviour of a vector of observables $y_t$ over time [2] units $t = 1, 2, ....$ The history of the sequence $y_t$ at time $t$ is given by $Y_t = \{y_s\}_{s=1}^{t}$. A 'model' is a corresponding sequence of probability density functions:

$$\pi_t(y_t|Y_{t-1}, \theta), \tag{3.5}$$

where $\theta$ is, say, $k \times 1$ vector of unknown parameters, $\theta \in \Theta \subseteq R^k$ and $t = 1, ....$ We will use $\pi(.)$ to denote both the distribution function and the density.

The probability density function of $Y_T$ conditional on the model and parameter vector $\theta$, is given by:

$$\pi(Y_T|\theta) = \prod_{t=1}^{T} \pi_t(y_t|Y_{t-1}, \theta). \tag{3.6}$$

The likelihood function is any function $L(\theta; Y_T) \propto \pi(Y_T|\theta)$.

If the model specifies that the $y_t$ are independent and identically distributed, then it holds that,

$$\pi_t(y_t|Y_{t-1}, \theta) = \pi_t(y_t|\theta), \tag{3.7}$$

$$\pi(Y_T|\theta) = \prod_{t=1}^{T} \pi_t(y_t|\theta). \tag{3.8}$$

The objective of Bayesian inference is to make useful inference about $\theta$ using the posterior distribution $\pi(\theta|Y_T)$. Computations for Bayes inference often reduce to computing expectations that in general can be expressed as follows:

$$E[h(\theta)|Y_T], \tag{3.9}$$

where $h(\theta)$ is a 'function of interest'. Broadly speaking, this function of interest can be:

[2] Subindex $t$ may pertain to cross sections, to time series, or both, but time series models and language are used here for specificity. Likewise, it is assumed that $y_t$ is continuously distributed for specificity and brevity.

- a parameter or a function of parameters,
- $h(\theta) = Lo(a_1, \theta) - Lo(a_2, \theta)$ in which $Lo(a, \theta)$ is the loss function associated to taking action $a$ with parameter vector $\theta$.
- $h(\theta) = \chi_{\Theta_0}(\theta)$ which arises when a hypothesis restricts $\theta$ to a set $\Theta_0$.[3]
- a predictive density, $h(\theta) = E[g(y^*)|Y_T, \theta]$, where $y^* = (y_{T+1}, ..., y_{T+l})$.

Models and functions of interest for which equation (3.9) can be analytically evaluated are quite limited.

The specification of the model in equation (3.8) is completed with a 'prior density' $\pi(\theta)$. It can be shown that given equation (3.8) and a density $\pi(Y_T)$ (i.e., a density for the data unconditional on $\theta$) a prior density must exist (Bernardo and Smith, 1996).

By Bayes Theorem the 'posterior density' of $\theta$ is given by:

$$\pi(\theta|Y_T) = \frac{\pi(Y_T|\theta)\pi(\theta)}{\pi(Y_T)} \tag{3.10}$$

$$\propto \pi(Y_T|\theta)\pi(\theta) \tag{3.11}$$

$$\propto L(\theta; Y_T)\pi(\theta). \tag{3.12}$$

Bayesian inference requires the evaluation of posterior expectations of functions of $\theta$. That is,

$$E[h(\theta)|Y_T] = \int_\Theta h(\theta)\pi(\theta|Y_T)d\theta = \frac{\int_\Theta h(\theta)L(\theta; Y_T)\pi(\theta)d\theta}{\int_\Theta L(\theta; Y_T)\pi(\theta)d\theta}. \tag{3.13}$$

In most applications, analytic evaluation of $E[h(\theta)|Y_T]$ is impossible. Alternative approaches are needed like Monte Carlo integration.

[3] $\chi(.)$ is the characteristic function. $\chi_s(z) = 1$ if $z \in S$, $\chi_s(z) = 0$ if $z \notin S$.

### 3.4.3 Monte Carlo Integration and Markov Chains

In this section a brief overview of some elementary Monte Carlo integration and Markov chains theory that will be needed to understand the following section will be covered. The material presented here can be found, for instance, in Meyn and Tweedie (1993). For more specific treatment of Markov chain theory relevant to the study of Markov chain Monte Carlo the reader is referred to Tierney (1994), Tierney (1996), Mengersen and Tweedie (1996) and Roberts and Tweedie (1996).

Let $x$ be a vector comprising the unknown model parameters, latent variables, missing values and/or future data with posterior distribution $\pi(.)$. As shown in the previous section, the task of Bayesian inference is to evaluate the posterior expectation:

$$E[h(x)] = \frac{\int h(x)\pi(x)\nu(dx)}{\int \pi(x)\nu(dx)}, \tag{3.14}$$

where $\pi$ is the density of $\pi$ with respect to the measure $\nu$ (hence, $\pi(dy) = \pi(y)\nu(dy)$) and $h(.)$ is the function of interest.

Monte Carlo integration evaluates $E[h(x)]$ by drawing samples $\{x_t, t = 1, ..., n\}$ from $\pi(.)$ and then approximating the expectation by, for example,

$$E[h(x)] \approx \frac{1}{n}\sum_{t=1}^{n} h(x_t). \tag{3.15}$$

In general, drawing samples $\{x_t\}$ independently from $\pi(.)$ is not feasible and, therefore, laws of large numbers that ensure that the approximation can be as accurate as desired by increasing the sample size $n$ are not applicable. However, the $\{x_t\}$ need not necessarily be independent. In fact, the $\{x_t\}$ can be generated by any process which, loosely speaking, draws samples throughout the support of $\pi(.)$ in the correct proportions. One way of doing this is through a Markov chain having $\pi(.)$ as its stationary distribution.

A Markov chain is a sequence of random variables (generally vectors), $x_0, x_1, x_2, ...$, such that at each time $t \geq 0$, the evolution of the chain on a space $\Omega \subseteq R^p$ is governed by the 'transition kernel':

$$\begin{aligned} P(x, A) &= P(x_{t+1} \in A | x_t = x, x_s, s < t) \\ &= P(x_{t+1} \in A | x_t = x), \end{aligned}$$

where $x \in \Omega, A \subset \Omega$.

The assumption that the probability distribution of the next state in the sequence, given the current and the past states, depends only on the current state is the 'Markov property'.

That is, given $x_t$, the next state $x_{t+1}$ does not depend further on the history of the chain $\{x_0, x_1, ..., x_{t-1}\}$. It will be supposed that the chain is time-homogeneous (i.e., $P(.|.)$ does not depend on $t$).

The $t$th step ahead transition kernel is consequently given by:

$$P(x_{t+m} \in A | x_s, s < m; x_m = x) = P^t(x, A); \tag{3.16}$$

that is, $P^t(x, A)$ denotes the probability that a chain at $x$ will be in the set $A$ after $t$ steps or transitions. The independence of $P^t$ on the values of $x_s$, $s < m$, is the Markov property, and the independence of $P^t$ and $m$ is the time-homogeneity property.

Under certain conditions, which will be briefly discussed later, it can be shown that the $t$th iterate of the transition kernel converges (as $t \to \infty$) to the invariant or stationary distribution, say $\pi$, that satisfies:

$$\pi(dy) = \int_\Omega P(x, dy)\pi(x)\nu(dx). \tag{3.17}$$

The invariance condition states that if $x_t$ is a sample from $\pi$ then all subsequent samples will be from that distribution. Thus, after a sufficiently long 'burn-in' of say $m$ iterations, points $\{x_t; t = m + 1, ...\}$ will be dependent samples approximately from $\pi$.

A chain is said to be 'reversible' if the transition kernel $P(x, y)$ satisfies 'detailed balance':

$$f(x)P(x, y) = f(y)P(y, x), \tag{3.18}$$

for a density, say, $f(.)$. If reversibility holds, it can be shown that $f(.) = \pi(.)$. Hence, a reversible chain has $\pi$ as its invariant distribution.

A Markov chain is said to be '$\pi$ -irreducible' if, for every $x \in \Omega$,

$$\pi(A) > 0 \Rightarrow P(x_t \in A | x_0 = x) > 0, \tag{3.19}$$

for some $t \geq 1$.

A third important property of a chain is 'aperiodicity', which ensures that the chain does not cycle through a finite number of sets.

Now the following (ergodicity) result, which forms the basis for MCMC methods, can be stated.

**Proposition 3.1.** *If $P(.,.)$ is $\pi$-irreducible and has invariant distribution $\pi$, then $\pi$ is the unique invariant distribution of $P(.,.)$. Then for $\pi$-almost every $x \in \Omega$ and all sets $A$, we have the following:*

- *For all $\pi$-integrable real-valued functions $h$,*
  $\frac{1}{t}\sum_{i=1}^{t} h(x_i) \rightarrow \int h(x)\pi(x)\nu(dx)$ *as* $t \rightarrow \infty$, *a.s.*

*If $P(.,.)$ is also aperiodic, then it holds,*

- $|P^t(x, A) - \pi(A)| \rightarrow 0$ *as* $t \rightarrow \infty$.

The first part of Proposition 3.1 tells us that averages of functions evaluated at sample values (*ergodic averages*) converge (as $t \rightarrow \infty$, almost surely) to their expected value under the target density. Later, sufficient conditions for '$\pi$-irreducibility' and aperiodicity will be presented.

The second part states that the probability density of the $t$th iterate of the Markov chain is, for large $t$, very close to its unique, invariant density. Hence, if drawings are made from $P^t(x, dy)$, then for large $t$ the probability distribution of the drawings is the invariant distribution, regardless of the initial values.

### 3.4.4 Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) methods have a history in mathematical physics dating back to the algorithm of Metropolis *et al* (1953). This method was generalised by Hastings (1970), who focused on statistical problems, and was further explored by Peskun (1973). A version particularly suited to image reconstruction and spatial statistics was then introduced by Geman and Geman (1984) and it was consequently shown to have great potential for Bayesian computation by Gelfand and Smith (1990).

Since 1990 applications of MCMC methods have grown rapidly, specially for Bayesian inference. New generic techniques and the mathematical properties of different methods continue to attract a huge research effort. For an introduction and overview of the theory implementation and applications of MCMC techniques see, for instance, the book by Gilks *et al* (1996).

The idea behind most approaches to construct a MCMC method is to generate a Markov chain that satisfies detailed balance (equation 3.18) and that, therefore, is reversible. This is a easy-to-work-with sufficient condition that guarantees that the target distribution is going to be the invariant distribution of the chain.

We are going to concentrate on the two most widely used MCMC methods, the Gibbs sampler and the Metropolis-Hastings algorithm.

### 3.4.5 Gibbs sampler

The Gibbs sampler is the most widely applied MCMC tool after the work of Geman and Geman (1984) which was popularised in statistics by the articles by Gelfand and Smith (1990) and Gelfand *et al.* (1990).

The value of this algorithm arises from the fact that in many applications the full conditional densities take convenient forms to be simulated even though the target density is intractable.

For Bayesian DAG models in particular, the joint distribution of the data and parameters is a product of many terms, each involving only a subset of the parameters. For such models, the full conditional distribution for any given parameter can be constructed from those few terms of the joint distribution which depend on it (Lauritzen, 1994, Spiegelhalter *et al.*, 1995).

Let $\pi(x) = \pi(x^1, ...x^k)$, $x \in R^n$, denote a joint density, and let $\pi(x^s|x^{-s})$ denote the induced full conditional densities for each of the components $x^s$, given values of the other components $x^{-s} = (x^l;\ l \neq s)$, $s = 1, ..., k \leq n$. Full conditional distribution densities are derived from the joint distribution:

$$\pi(x^s|x^{-s}) = \frac{\pi(x^s, x^{-s})}{\int \pi(x^s, x^{-s})dx^s}. \tag{3.20}$$

We start with a partition or 'blocking' of the random vector $x = (x^1, ...x,^k)$. One important consideration in choosing the level at which the components for the conditionals are chosen is the correlation structure of $\pi(x)$. If highly correlated scalar components are treated individually, there could be slow convergence of the chain to its stationary distribution density as a result of very little movement at each conditional random variate generation step. If, however, correlated scalars are blocked together to form a subvector component this problem is avoided, but at the expense of having to perform a draw from

a multivariate conditional distribution (see, for instance, Roberts and Sahu, 1997).

A systematic form of the Gibbs algorithm proceeds as follows. Start by picking arbitrary starting values, say $x_0 = (x_0^1, ..., x_0^k)$. Then successively make random drawings from the full conditional distributions $\pi(x^s|x^{-s})$, $s = 1, ..., k$, as follows:

$$\begin{aligned} x_1^1 &\text{ from } \pi(x^1|x_0^{-1}); \\ x_1^2 &\text{ from } \pi(x^2|x_1^1, x_0^3, ..., x_0^k); \\ x_1^3 &\text{ from } \pi(x^3|x_1^1, x_1^2, x_0^4, ..., x_0^k); \\ &\vdots \\ x_1^k &\text{ from } \pi(x^k|x_1^{-k}). \end{aligned}$$

This completes a transition from $x_0 = (x_0^1, ..., x_0^k)$ to $x_1 = (x_1^1, ..., x_1^k)$. Iterating this cycle of random variate generation from each of the full conditional distributions in turn produces a sequence $x_0, x_1, ..., x_t, ...$ which is a realisation of a Markov chain, with transition probability from $x_t = x$ to $x_{t+1} = y$ given by:

$$p_G(x, y) = \prod_{s=1}^{k} \pi(y^s|y^1, ..., y^{s-1}, x^{s+1}, ..., x^k). \tag{3.21}$$

Suppose that $\nu$ is the Lebesgue measure [4]; then,

$$\begin{aligned} \int p_G(x, y)\pi(x)dx \;=\; & \int ... \int \prod_{s=1}^{k} \frac{\pi(y^s|y^1, ..., y^{s-1})\pi(x^{s+1}, ..., x^k|y^1, ..., y^s)}{\pi(x^{s+1}, ..., x^k)dx} \\ & \times\; \pi(x^1|x^2, ..., x^k)\pi(x^2, ..., x^k)dx, \end{aligned}$$

by applying Bayes theorem to each term in the transition kernel, letting $y^0$ denote the empty set, and writing $\pi(x)$ as $\pi(x^1|x^2, ..., x^k)\pi(x^2, ..., x^k)$. The calculation is completed by noting that,

[4] On the real line $(R, B(R))$ Lebesgue measure $\nu$ is a positive measure defined for intervals $(a, b]$ by $\nu(a, b] = b - a$, and for the other sets in $B(R)$ by an obvious extension technique. Lebesgue measure on higher dimensional Euclidean space $R^p$ is constructed similarly using the area of rectangles as a basic definition.

- the terms $\pi(y^s|y^1,...,y^{s-1})$ are independent of $x$, so they factor out as $\prod_{s=1}^{k}(y^s|y^1,...,y^{s-1})$ to give $\pi(y)$;
- the integral over $x^1$ is 1;
- the term $\pi(x^2,...,x^k)$ cancels for $s=1$; and
- cancellation by telescoping occurs because the numerator element in term $s-1$ is $\pi(x^{s+1},...,s^k|y^1,...,y^{s-1})$ after the requisite integration over $x^k$, which cancels with the denominator in term $s$.

A convenient set of sufficient conditions which ensures that the Markov chain generated by the Gibbs sampler satisfies the conditions in *Proposition 3.1* is due to Roberts and Smith (1994).

**Proposition 3.2.** *Suppose that*

- $\pi(x) > 0$ *implies there exists an open neighborhood* $N_x$ *containing* $x$ *and* $\epsilon > 0$ *such that, for all* $y \in N_x$, $\pi(y) \geq \epsilon > 0$;
- $\int \pi(x)dx^s$ *is bounded for all* $s$ *and* $y \in N_x$ *in an open neighborhood of* $x$; *and*
- *the support of* $x$ *is arc-connected.*

*Then* $p_G(x,y)$ *satisfies the conditions of Proposition 3.1.*

The conditions ensure that each full conditional density is well defined and that the support of the density is not separated into disjoint regions so that once the chain moves into one such region it never leaves it. Although these are only sufficient conditions for the convergence of the Gibbs sampler, the conditions are extremely weak and are satisfied in most applications.

In general, four steps are required to implement the Gibbs sampling:

- starting values must be provided for all unobserved nodes;
- full conditional distributions for each unobserved node must be constructed and methods for sampling from them decided upon;
- the output must be monitored to decide on the length of the *burn-in* and the total run length, or perhaps to identify whether a more computationally efficient parameterisation or MCMC algorithm is required;
- summary statistics for quantities of interest must be calculated and examined from the output, for inference about the true values of the unobserved nodes .

### 3.4.6 Metropolis-Hastings

The Metropolis-Hastings algorithm is due to the generalisation that Hastings (1970) made of the algorithm introduced by Metropolis *et al.* (1953).

Following with the notation presented previously, the Metropolis-Hastings algorithm begins with a candidate new value, or 'proposal', $y$ by drawing $y \in R^n$, from an arbitrary transition probability density function $q(y;x)$ which generates transitions for a discreet time Markov chain evolving on $R^n$. We are going to update $x$ in a single block without loss of generality.

The algorithm will actually accept the proposal with probability,

$$\alpha(x,y) = \begin{cases} \min\{1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}\} & \text{if } \pi(x)q(x,y) > 0 \\ 1 & \text{if } \pi(x)q(x,y) = 0 \end{cases} . \tag{3.22}$$

This defines a Markov chain with a transition probability from $x_t = x$ to $x_{t+1} = y$ given by:

$$p_{\text{MH}} = \begin{cases} q(x,y)\alpha(x,y) & \text{if } y \neq x \\ \int q(x,y)[1-\alpha(x,y)]dy & \text{if } y = x \end{cases} . \tag{3.23}$$

Note that the Gibbs sampler is the special case of the Metropolis-Hastings algorithm where the proposal density $q(y;x)$ is the full conditional density $\pi(y|x)$, so that the proposal is accepted with probability one.

The sequence $x_t$ generated by the Metropolis-Hastings algorithm satisfies detailed balance (see equation 3.18) by construction of $\alpha$ and, therefore, its equilibrium distribution is $\pi$.

In implementing the Metropolis-Hastings algorithm, the transition probability density function must share two important properties: it must be possible to generate $y$ from $q(.)$ easily and the acceptance rate should not be so low that the time required to generate a sufficient number of different $x_t$ is too great (see, for instance, Besag *et al*, 1995 and Gelman *et al*, 1996).

The convergence properties of the Metropolis-Hastings algorithm have been extensively studied following the work by Roberts and Smith (1994) and Mengersen and Tweedie (1996) and sufficient conditions found by Tierney (1994).

## 3.5 Performance of MCMC methods

MCMC methodology relies on the fact that the $t$-step transition kernel of the algorithms converges to the target density as $t \to \infty$. When implementing any of the MCMC algorithms, there are innumerable strategies affecting the 'efficiency' of the algorithm. Efficiency consists, broadly speaking, of reducing the number of *burn in* iterations (*i.e.*, the initial iterations which are considered not to arise from the target density) and the amount of required calculations at each iteration.

Consequently, there are different techniques and recipes to construct appropriate algorithms in different scenarios that have been, and are being, proposed

in the literature. For an overview of different implementation and practical issues the reader is referred to the book by Gilks *et al* (1996). In what follows, just a summary of few of these topics, specially those concerning convergence of the chains, will be presented.

Convergence diagnostics are fundamental in order to establish that the sample obtained by running any algorithm constitutes a sample from the distribution of interest. Much research is being made for diagnosing convergence of Gibbs samplers and other Markov chain Monte Carlo algorithms (see, for instance, Brooks and Roberts (1995), Cowles and Carlin (1996), and Brooks and Roberts (1998)). No particular method of assessing convergence is universally accepted as being foolproof and it is recommended using a combination of diagnostics in addition to visual inspection of the trace plots in order to be able to assess convergence with a reasonable degree of confidence.

The literature has suggested two methods for generating a sample from a MCMC algorithm: the single chain and the multiple chain. In the multiple-chain method, the value at the end of a 'burn-in' period of , say, $n$ drawings is taken as a draw from the target distribution and the process is repeated with a new starting value. This method is somehow considered wasteful because it generates an independent sample at the cost of discarding $n$ drawings in each run. The multiple-run method has been superseded for the most part by the single-run method. Even though this scheme usually introduces strong correlation between parameter values at successive iterations, the correlation often dissipates quickly for moderate $n$ .

Because the length of the 'burn-in' period seems to be model – and data – dependent, the question of convergence requires considerable care. If the target density being simulated is 'well behaved' (as it is in many standard applications), then the simulated Markov chain usually mixes rapidly and the serial correlations die out quickly. But with weak identifiability of the parameters

and/or multiple modes the chain can be poorly behaved and slow to produce numerically accurate results. Many proposals have been made to shed light on these problems.

One class of approaches (Ritter and Tanner, 1992; Gelman and Rubin, 1992; Geweke, 1992; Zellner and Min, 1995 and Brooks and Gelman, 1998) attempts to analyse the output to determine whether or not the chain has converged.

The Gelman and Rubin (Gelman and Rubin, 1992) and the Brooks, Gelman and Rubin approach (Brooks and Gelman, 1998), which is based on multiple-chain sampling from dispersed starting values, compares the within and between variation in the sampled values. The Ritter and Tanner approach (Ritter and Tanner, 1992), which requires a single run, monitors the ratio of the target density (up to a normalising constant) and the current estimate of the target density; stability of the ratio indicates that the chain has converged. Geweke's (Geweke, 1992) convergence diagnostic, which is based on standard time series methods, is appropriate for the analysis of individual chains when convergence of the *mean* of some function of the sampled variables is of interest.

Another type of approach (Raftery and Lewis, 1992, and Polson, 1994) attempts to produce estimates of both the 'burn-in' time prior to sampling by analysing the rate of convergence of the Markov chain to the target density and the run-lengths needed to accurately estimate quantiles of functions of the variables of interest.

In addition to different convergence diagnostics, a preliminary visual analysis of the obtained chains is strongly recommended. A rapid mixing chain will generally present rapid fluctuations over the sample space. High autocorrelations within chains and substantial correlation between variables usually indicate slow mixing and, therefore, slow convergence. This will be characterised by

plots of sample traces which 'snake' slowly up and down. Reparameterisation of the model variables can sometimes lead to less correlated chains. space.

## 3.6 Discussion

In this chapter a general Bayesian population pharmacokinetic hierarchical model, accommodating different pharmacokinetic models between and within individuals, has been built. The DAG representation has allowed to specify the full joint distribution of a rather complicated model in an intuitive and simple manner from local components. This factorisation, in addition, facilitates the implementation of MCMC tools that overcome the historical constrain that Bayesian inference has experienced: difficulties computing the posterior distribution.

# Chapter 4

# A Case Study: Cyclosporin

After reviewing the main pharmacokinetic models and, we believe, the most adequate inference framework, we are ready to handle real data. A fully Bayesian three-stage hierarchical model is going to be fitted to a real data set. The most suitable pharmacokinetic model needs first to be chosen following visual inspection of the data and literature recommendations. Once the structural model is built, probabilistic assumptions will be made and estimation performed.

## 4.1 Introduction

Cyclosporin is in routine use in organ transplantation processes as an inmunosuppressive medication. Fahr (1993) constitutes an excellent summary of all the research and literature survey of the clinical pharmacokinetics of cyclosporin. Even though it is prescribed by the medical profession on a daily basis, its pharmacokinetics are of great importance for a specialised group of clinical pharmacologist as standard dosing regimes of cyclosporin do not work because its therapeutic range is narrow and the variability of its pharmacokinetics are large between different patients and within individual patients. A narrow therapeutic range, with the consequent high risk of lethal poisoning,

with a poor pharmacokinetic understanding originates conservative doses that can be clinically uneffective.

There are two approaches in designing dose schedules:

- Administration of standard subtherapeutic, but safe, doses in conjunction with therapeutic doses of other better understood immunosuppressive medications.
- Individualised administration after monitoring cyclosporin blood concentrations that requires a reasonable understanding of the individual specific pharmacokinetic characteristics.

The latter approach has got an increasing statistical interest and will be introduced in this chapter.

## 4.2 Pharmacokinetic Variability

The pharmacokinetics of cyclosporin, as mentioned before, are very variable both between and within patients. The sites of action are rather unknown and extremely complex and, even though attempts have been done in order to quantify the drug concentration in different sites of action, they are all rather time-consuming and lack specificity for cyclosporin. Currently, blood concentration monitoring constitutes the only clinically useful feedback information on which dosing decisions can be based.

The first source of pharmacokinetic variability of any drug is related to the error associated with the method chosen to measure drug concentrations in blood. Different techniques show diverse degrees of accuracy and precision. This source of variability may diminish in the near future if the recommendations of recent consensus documents to monitor blood concentrations are

followed. Nevertheless, the pharmacokientic parameters are highly variable and depend on many individual factors.

Clinical studies show that the bioavailability of cyclosporin varies dramatically between and within individuals depending on a large number of factors such as the mixing of the oral formulation with the drink, individual physiological characteristics, interactions with other drugs, etc.

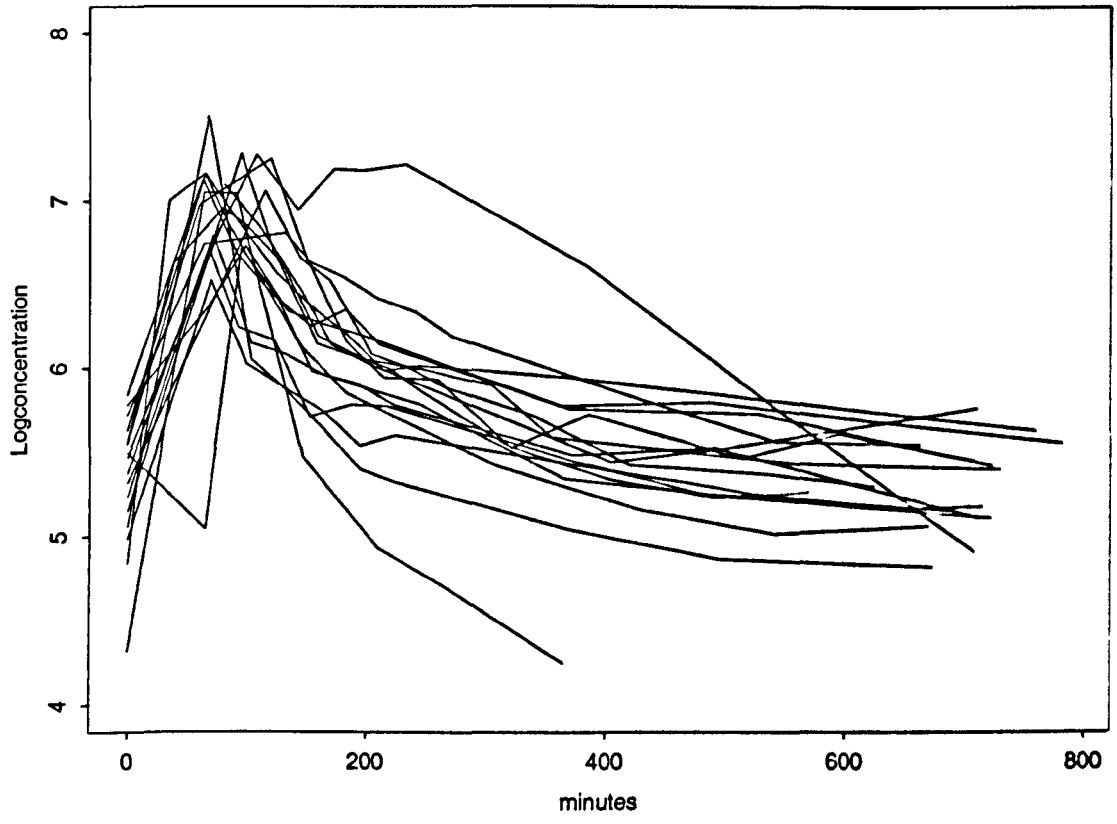
Cyclosporin is not uniformly distributed throughout the body, and the literature suggests that it is extensively bounded within the blood compartment and even more to tissues. Nonetheless, little is known with accuracy about the distribution process of cyclosporin due to its complexity and, again, variability. Little work has been done up to the present on the pharmacokinetics of cyclosporin following the population approach using either classical or Bayesian tools (Rui *et al.*, 1995, Charpiat *et al.*, 1998, Parke and Charles, 1998, Jensen and Dalhoff, 2000).

Elimination occurs principally by metabolism in the liver, and any factor affecting the liver will influence the extent and rate of elimination.

## 4.3 Data

The data set we are concerned with is presented in figure 4.1. This data set was provided by Dr. Stephen Lowis from the Bristol Royal Hospital for Sick Children. A group of seventeen patients received different doses of cyclosporin intravenously. Logarithmic concentration level observations are plotted against time. The doses were administered by infusion during approximately an hour, varying between patients. After and during administration, blood samples were collected for each patient at different times and blood cyclosporin concentration levels determined following the same technique to monitor blood

Figure 4.1: Cyclosporin Blood Concentration. Last dose of the treatment is given at time 0. The logarithm of observed concentration levels for 17 patients are plotted against time.




concentrations. They all had previous doses, and that is why the starting concentration levels are in all cases positive.

Figure 4.1 shows how concentration levels increase abruptly during the infusion of the substance. Drug concentration levels reach a peak, shortly after the infusion is terminated , approximately in 60 minutes. After reaching the peak, drug distribution and elimination impose an exponential decay of blood logarithmic drug levels, suggesting that a two compartmental model might be necessary and sufficient in order to explain the drug blood concentration

behaviour. The first rapid decay is probably originated by the distribution of the substance from the circulatory system to different organ and tissues in addition to the fraction of drug that is being eliminated from the body. Then follows a somehow more moderate decrease due to the combination of two processes: part of the drug is being expelled from the body and some is being redistributed into the circulatory system.

The data set looks at first sight fairly homogeneous with the possible exception of a couple of patients. We aim to construct a model flexible enough to accommodate this kind of irregularities between individuals.

## 4.4 The Pharmacokinetic Model

The logarithmic concentration - time plots in Figure 4.1 suggest that a two compartment model should be first tried to explain the data we are concerned with. Blood samples have been taken after the beginning of the administration of the last dose, at time zero, of a treatment that so far comprises 20 previous doses given at regular intervals of 8 hours.

Recall that (see section 2.5) for an intravenously administered infusion single dose, in a two-compartment model with first order distribution and elimination, the blood plasma drug concentration at any time is given by the following equation:

$$C^1(t) = A_1(e^{-a_1 t} - e^{-a_1 t_1'}) + A_2(e^{-a_2 t} - e^{-a_2 t_1'}), \tag{4.1}$$

where, defining $T_1^I$ as the length of infusion, $t_1' = t - T_1^I$ for $t > T_1^I$ and $t' = 0$ otherwise. The constants $A_1$ and $A_2$ are given by,

$$A_1 = \frac{D_1}{T_1^I V_1} \frac{(k_{21} - a_1)}{(a_1 - a_2)a_1} \quad , \tag{4.2}$$

$$A_2 = \frac{-D_1}{T_1^I V_1} \frac{(k_{21} - a_2)}{(a_1 - a_2)a_2} \quad , \tag{4.3}$$

where, $a_1$ and $a_2$ $(a_1 > a_2)$ are the roots of the quadratic equation: $r^2 + (k_{12} + k_{21} + k_{el})r + k_{21}k_{el} = 0$.

As it has been said above, all the patients have been given 20 previous doses, so that the pharmacokinetic equation from the beginning of the $21^{st}$ infusion for each individual is given by:

$$C(t) = \sum_{d=1}^{21} C^d(t). \tag{4.4}$$

## 4.5 The Probability Model

Let us denote the $j^{th}$ observed concentration for patient $i$ $(i = 1, ...K$, where $K$ is the number of patients) by $c_{ij}$ and the time at which the associated sample was collected by $t_{ij}$. The dosing history for patient $i$ at time $t_{ij}$ will be described by $H_{ij}$, and $\theta_i$ will denote the $p \times 1$ vector of pharmacokinetic parameters for individual $i$.

As discussed in chapter 3, pharmacokinetic models fit naturally into a three-stage hierarchical model (Steimer *et al.*, 1994, Wakefield *et al.*, 1994, Bennett *et al.*, 1995, Wakefield and Racine-Poon, 1995). In Figure 4.2. the 'directed acyclic graph' (DAG) (see section 3.3 for notation and conventions) representing the structural assumptions we will make below is shown.

At the first stage of the hierarchical model, the form of the probability distribution of each observed concentration level, $c_{ij}$, given the individual pharmacokinetic parameters, $\theta_i = (V, k_{12}, k_{21}, k_{el})$ and $\tau$ needs to be specified, where $\tau$, the inverse of the residual error variance, is assumed to be a normally distributed multiplicative error. Therefore, denoting by $C_{ij}$ the mean, unobservable, concentration level given by equation (4.4), it follows that,

$$[\ln(c_{ij})|\theta_i, \tau] = N(\ln(C_{ij}), \tau^{-1}), \tag{4.5}$$

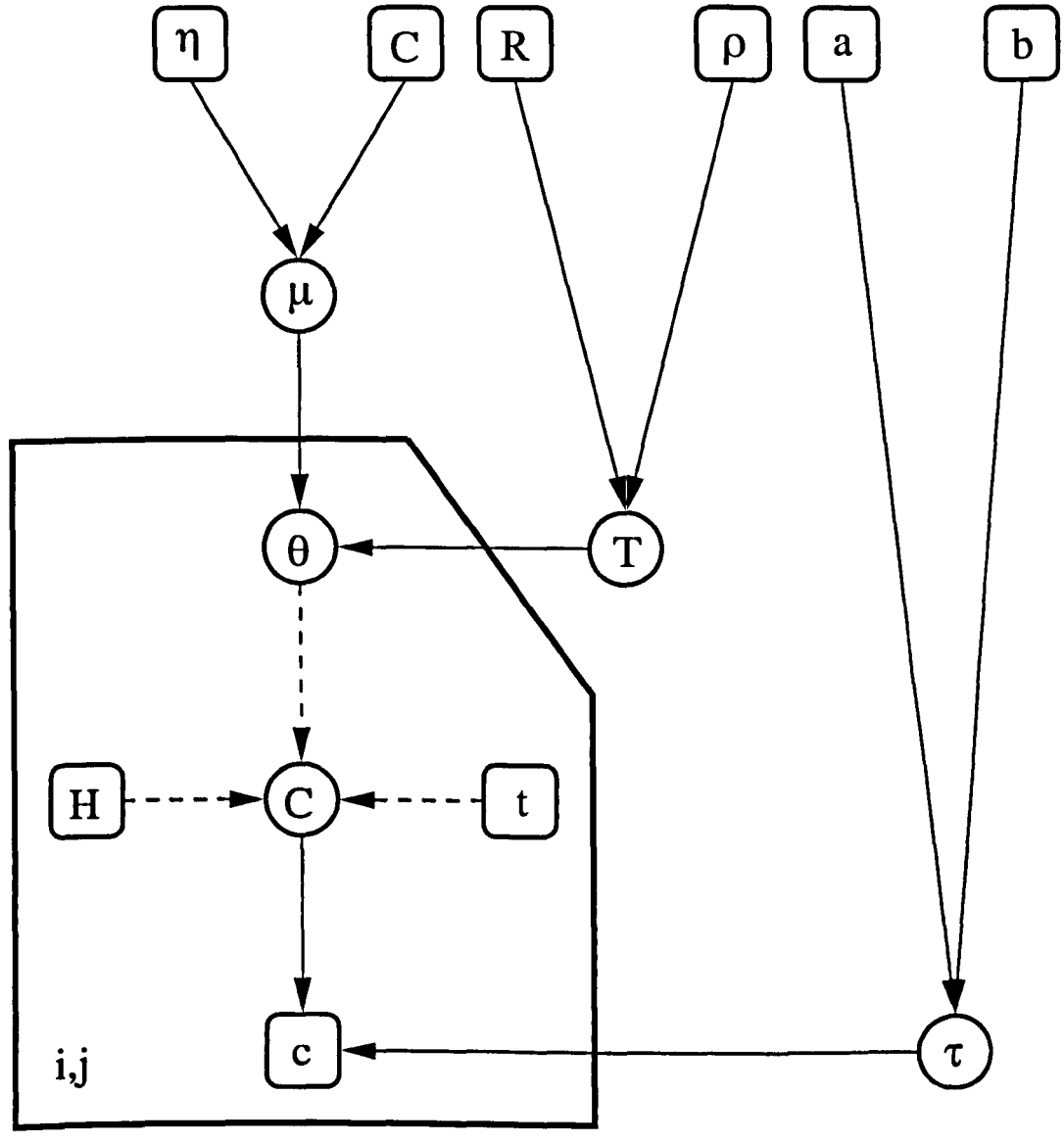



Figure 4.2: DAG representation for the Cyclosporin Model. Due to the lack of relevant covariate information about the patients, the structure of the model is more simple than in the general model presented in the previous chapter.

where we follow the notation of Gelfand and Smith (1990) for defining probability densities, i.e., joint, conditional and marginal densities of the random variables $x$ and $y$ are denoted by $[x, y]$, $[x|y]$ and $[x]$, $[y]$ repectively. A normal distribution with mean $a$ and precision parameter $b$ is as usual denoted by $N(a, b)$.

Exploiting the conditional independence property and denoting by $c$ the totality of the observed data and by $\Theta = \{\theta_1, ..., \theta_K\}$, we have that:

$$[\ln(c)|\Theta, \tau] = \prod_{i=1}^{K} \prod_{j=1}^{n_i} N(\ln(C_{ij}), \tau^{-1}). \tag{4.6}$$

At the second stage of the probabilistic model, assumptions regarding the distribution of each patient's pharmacokinetic parameters must be done [1]:

$$[\theta_i|\mu, \Gamma] = N_p(\mu, \Gamma^{-1}) \quad \text{for } i = 1, ..., K. \tag{4.7}$$

And, therefore, due to the conditional independence property, it follows that

$$[\Theta|\mu, \Gamma] = \prod_{i=1}^{K} N_p(\mu, \Gamma^{-1}), \tag{4.8}$$

where $N_p(A, B)$ denotes a $p$-dimensional multivariate Normal distribution with $p \times 1$ mean vector $A$ and $p \times p$ matrix of variance-covariance $B$. Here $\mu$ represents the mean population pharmacokinetic parameters and $\Gamma^{-1}$ the interindividual variability of the pharmacokinetic parameters.

The stochastic model is completed by the specification of the third stage, in which prior densities are assigned to the parameters $\theta$, $\Gamma$ and $\tau$. The distributions we choose are:

$$[\tau] = \text{Ga}(a, b), \quad [\mu] = N_q(\eta, C), \quad \text{and} \quad [\Gamma] = W_p(R, \rho). \tag{4.9}$$

where $W(A, b)$ denotes a Wishart distribution with $b$ degrees of freedom and scale matrix $A$ and $\text{G}(a, b)$ a Gamma distribution with parameters $a$ and $b$.

[1] There is no relevant covariate information recorded.

The last step is to assign values to the hyperparameters of the model, *i.e.* $a$, $b$, $\eta$, $C$, $R$ and $\rho$. The chosen value for both $a$ and $b$ is 0.001. Consequently, $\tau$ has a prior expected value of 1 and prior standard deviation of 1000. This prior is, therefore, proper but very vague as there is not any kind of prior information to include in the model about the real value of $\tau$.

To represent vague prior knowledge about $\mu$, we propose $\eta = (0, 0, 0, 0)$ and $C$ as a diagonal matrix taking values $10^{-4}$, again, a very uninformative, but still proper, distribution.

Lastly, as there is no available prior information about $\Gamma$, in order to have an uninformative prior distribution, we take the degrees of freedom for this distribution, $\rho$, to be as small as possible (*i.e.* 4, the rank of $\Gamma$). The scale matrix $R$ represents our prior guess of the order of magnitude of the covariance matrix. Note that except for cases with very few individuals, the choice of $C$ has little effect on the posterior estimate of $\Gamma_u$ (Lindley, 1970). We set the off-diagonal elements to zero, and the diagonal elements are set to 3.1 [2].

Once the distributional assumptions of this three-stage hierarchy are made, we are ready to construct the joint posterior distribution. Appealing to Bayes' theorem, the joint posterior is proportional to the product of the prior densities and the likelihood function (see section 3.4.2)

## 4.6 Estimation

The software **PKBUGS**(94) has been used to perform the estimation of the posterior distributions of the model parameters conditional on the given data. **PKBUGS** uses the Gibbs sampler (Geman and Geman, 1984) to explore the posterior distributions of the parameters $\theta$, $\mu$, $\tau$, and $\Gamma$. The components of

[2] We follow the recommendations given by the authors of *PKBUGS*.

Table 4.1: Summary: Marginal Posterior Distributions for the Population Pharmacokinetic Parameters.

| ln(Param) | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|
| $\mu_1 = C_L$ | -1.769 | 0.1277 | 0.003235 | -2.021 | -1.769 | -1.517 |
| $\mu_2 = Q$ | -1.073 | 0.2223 | 0.005898 | -1.492 | -1.08 | -0.6078 |
| $\mu_3 = V_1$ | 3.32 | 0.2312 | 0.00694 | 2.868 | 3.322 | 3.779 |
| $\mu_4 = V_2$ | 4.82 | 0.3463 | 0.01552 | 4.212 | 4.795 | 5.588 |

$\Theta$ are sampled using a Metropolis-Hastings algorithm (Metropolis *et al.*, 1953, and Hastings, 1970).

All stochastic parameters in the population pharmacokinetic model, *i.e.* $\Theta$, $\tau$, $\mu$ and $\Gamma$, must be given initial values. The values we choose are the prior means.

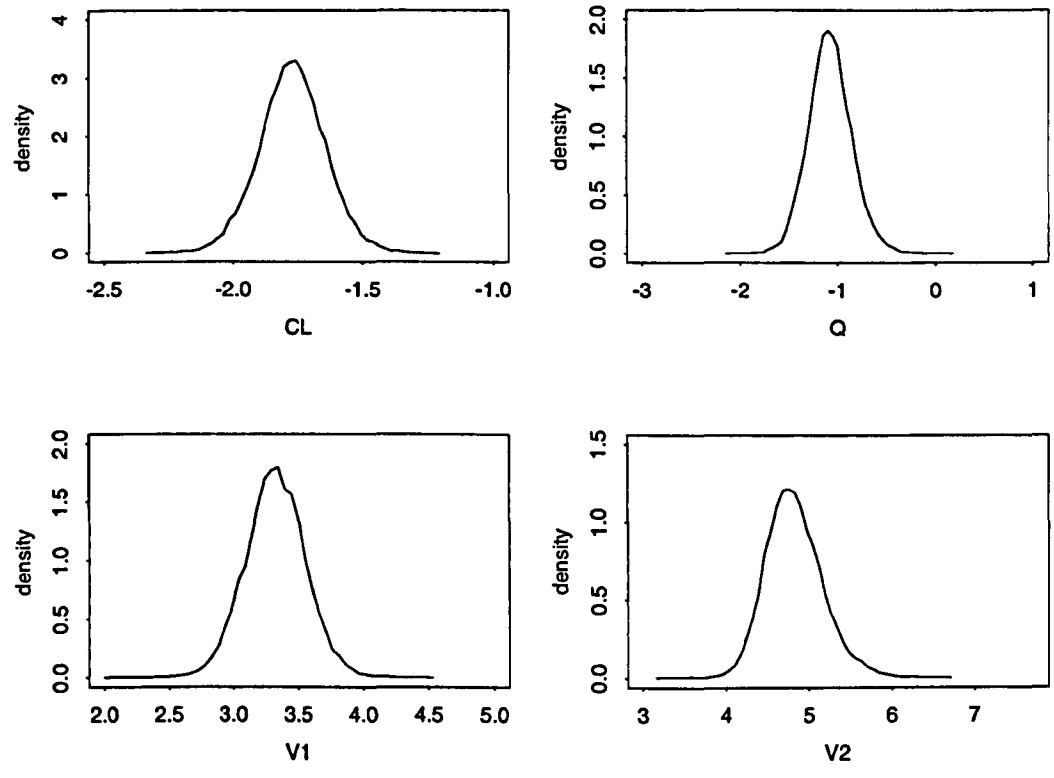
After several attempts, the first 20.000 iterations of the Gibbs sampler were discarded as 'burn-in', in order to, hopefully, guarantee that the stationary distribution was reached. Then, the chains were run for an additional 20.000 iterations past 'burn-in'. The computation took around 25 minutes.

In table 4.1, the estimation of the posterior marginal distributions of the population pharmacokinetic parameters is summarised. Here $C_L$ represents the clearance, $Q$ denotes the distributional clearance, $V_1$ and $V_2$ the volume of compartments 1 and 2 respectively. All the variables are on a logarithmic scale. To get back the rate constants in equation (4.4) from these parameters, we apply:

$$C_L = k_{el}V_1, \quad Q = k_{12}V_1 = k_{21}V_2. \tag{4.10}$$

The second column on table 4.1, reports the expected value of the posterior marginal distribution of the corresponding parameter in the first column. The

Figure 4.3: Estimated posterior distributions of population parameters.




third column gives the standard deviation of the distribution. The *MC error* is the standard error of the estimate of the expected value in column 2; it is estimated by considering both the sample size and the degree of autocorrelation in the sample. The last three columns give the quantiles and the median of the distributions.

The standard deviations and the quantiles of the posterior distributions of the population parameters suggest that these distributions are reasonably symmetric and concentrated around their mean values, as expected. Therefore, and taking in account the small *MC errors*, these expected values are statistically *good* point estimators of the real values. The estimated posterior

distributions of population parameters, $\theta$, calculated from the samples of each of the variables are shown in figure 4.3. PKBUGS carries out the Kernel density estimation using the S-Plus 'density' function. This leads to a smooth estimate that could hide local features of the density.

Individual pharmacokinetic parameters' posterior distributions are summarised in Appendix A.

Substituting posterior expected values of $\mu$ in equation (4.4), estimations of observed concentration levels are calculated. The results that we obtain are represented in figures 4.4 – 4.6. Data points are represented by dots. *Individual curves* are obtained by substituting the individual parameters' posterior distributions' means, together with the corresponding doses and infusion lengths, in model 4.4. These are represented by solid curves. To get an insight into the population variability, we can construct a curve (the dotted curves in figures 4.4 – 4.6) where individual variability is removed by introducing the means of the posterior distributions of the population parameters together with individual doses and infusion lengths in model 4.4. We refer to these as 'population curves'.

In general terms, the individual curves are encouraging and seem to reflect that the chosen two compartment model captures in a simplified fashion most of the pharmacokinetic processes that cyclosporin induces in the organism within this particular scenario (*e.g.* possible interactions with other drugs administered in the personalised treatments, the particular set of patients that enter in our study, etc.).

Differences between individual curves and the population counterparts, give information about the heterogeneity of the pharmacokinetic profiles of different patients. For example, if the population curve lies markedly under the individual curve, the concentration levels of the individual are higher than those

'expected'. This could be, for example, due to a relatively small body mass of the patient (information that is not available). It can be seen that the population curve differs substantially from the individual curve for approximately half of the patients. Therefore, our data set corroborates what has been implied by previous studies, *e.g.*, that the pharmacokinetics of cyclosporin show an important variability between individuals.

Let us point out that the second observation of Patient 3 seems to be somehow 'strange' as the concentration actually decreases during the intake of the drug. This could be due to the slow elimination of previous doses. Nevertheless, we believe that this observation has been measured with some sort of error. The consequences are obvious in figure 4.4 as it influences the estimated curve notably due to the small amount of observations. If our suspicion were corroborated, we would probably delete that observation in order to get better fittings for the rest of the observations.

There are relatively few blood samples for Patient 12, but the estimated curve explain pretty well the observations.

Blood concentrations for Patient 13 show a very particular profile and the use of a two compartment model without any covariate information seems not to be enough to estimate the concentration levels for this patient. Nevertheless, comparing the individual curve with the corresponding population one, it can be seen that the our model is flexible enough to capture substantial individual variability.

## 4.7 Convergence diagnostic

All the analysis presented above requires that the Gibbs sampler and the Metropolis-Hastings algorithm have effectively converged to their unique sta-

Figure 4.4: Patients 1-6, ordered from right to left, top to bottom. Log concentration - time. Dots represent observed data, solid lines are the individual curves, and dotted lines, the corresponding population curves. D = Dose. TI = Infusion Interval.

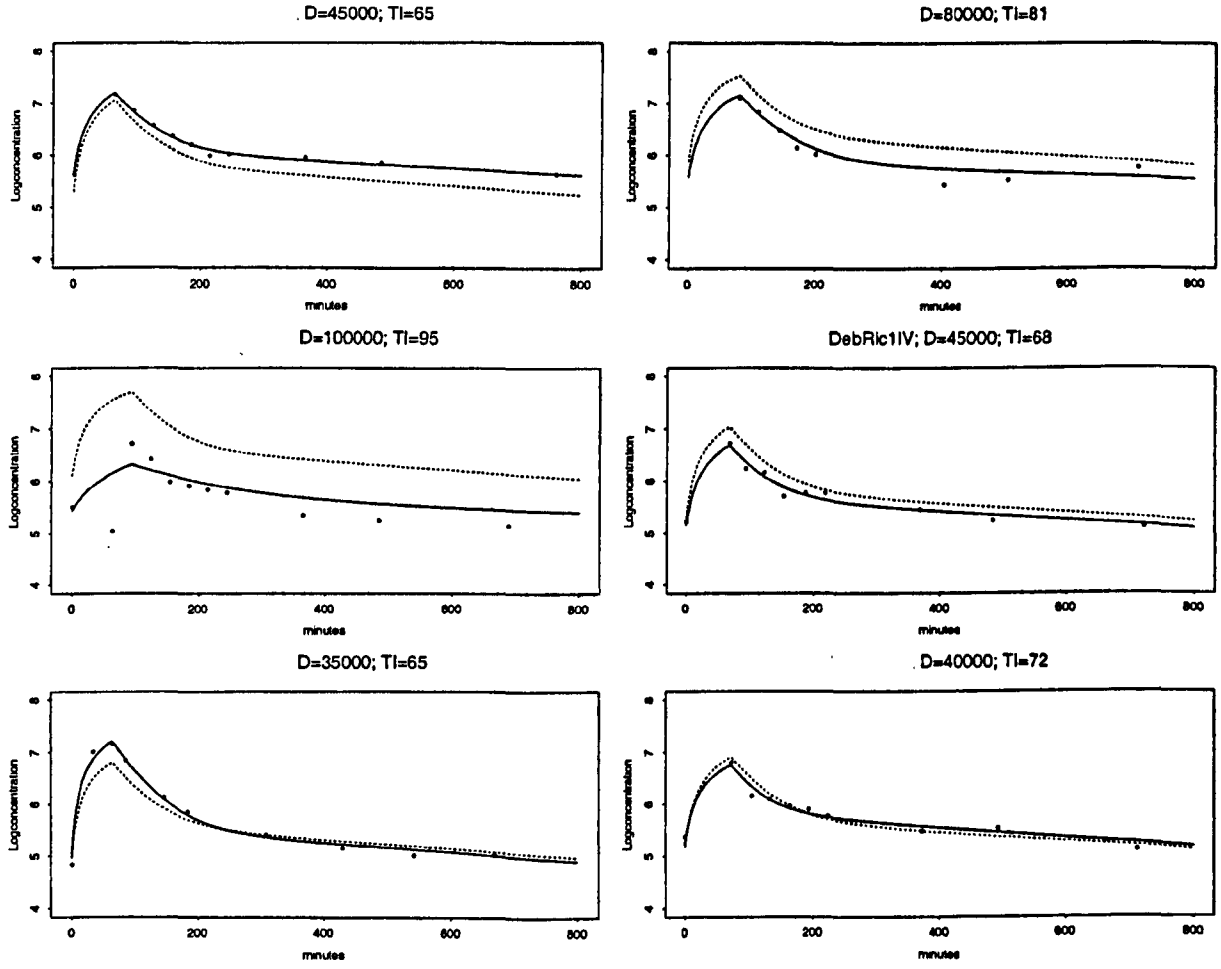


Figure 4.5: Patients 7-12, ordered from right to left, top to bottom. Log concentration - time. Dots represent observed data, solid lines are the individual curves, and dotted lines, the corresponding population curves. D = Dose. TI = Infusion Interval.

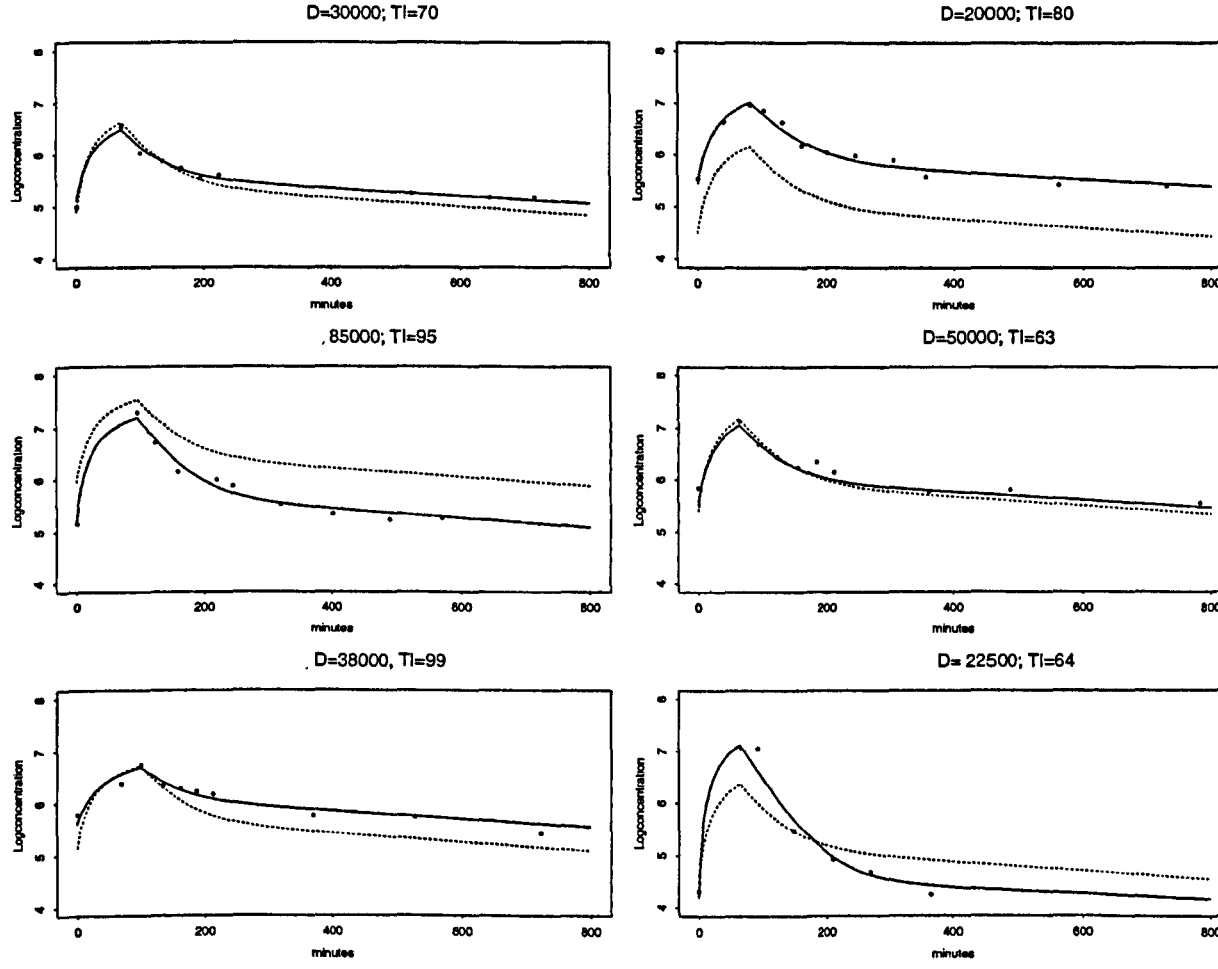


Figure 4.6: Patients 13-17, ordered from left to right, top to bottom. Log concentration - time. Dots represent observed data, solid lines are the individual curves, and dotted lines, the corresponding population curves. D = Dose. TI = Infusion Interval.

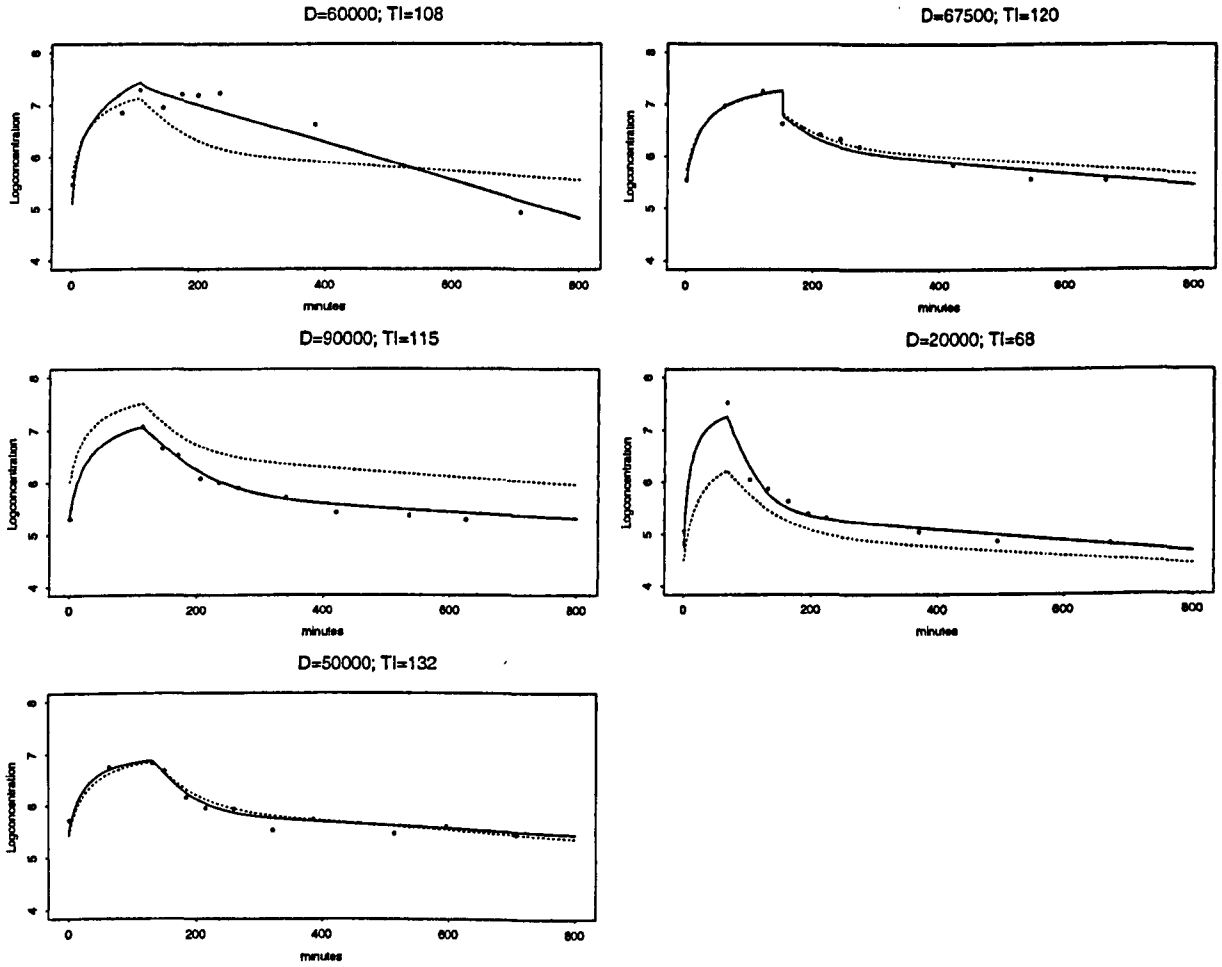
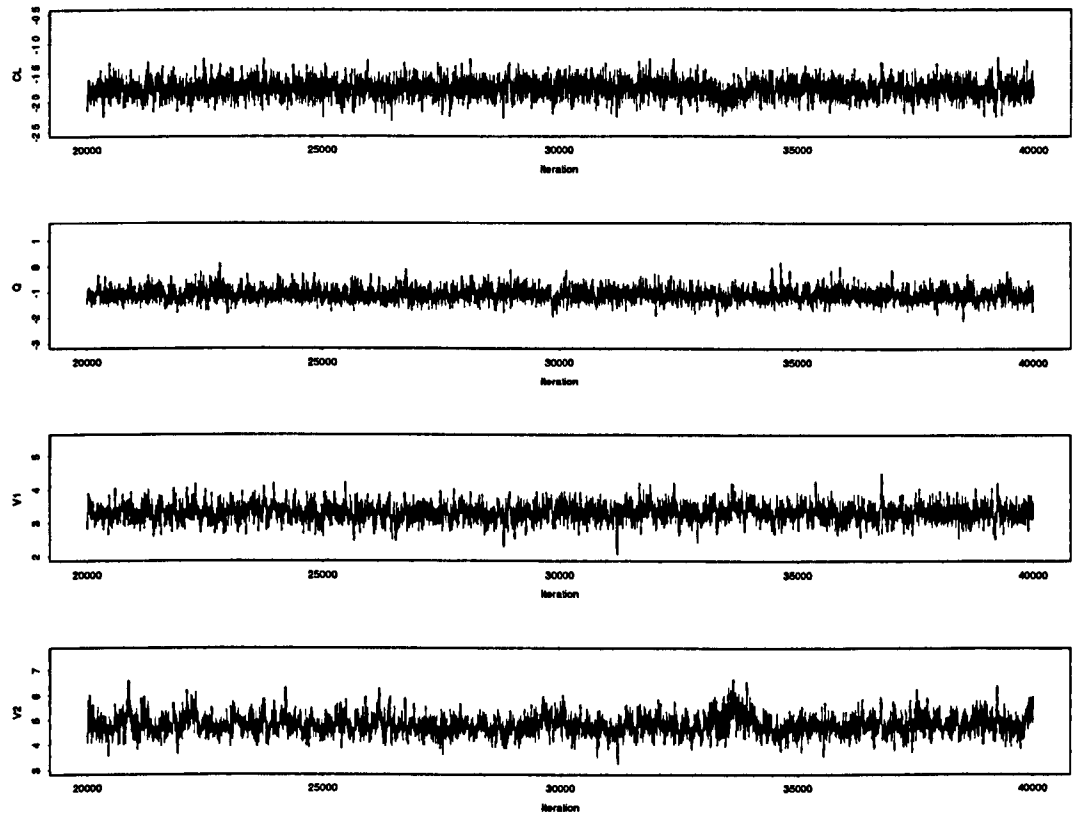


Figure 4.7: Traces of population mean parameters.




tionary or invariant distributions. Convergence diagnostics are fundamental in order to establish that the sample obtained by running the algorithms constitute a sample from the posterior distribution of interest.

As briefly discussed in chapter 3, much research is carried out for diagnosing convergence of Gibbs samplers and other Markov chain Monte Carlo sequences (see, for instance, Brooks and Roberts (1995)). No particular method of assessing convergence has universally been accepted as being superior in different scenarios.

The software PKBUGS does not include any tool for analysing convergence. In what follows, only a graphical inspection of the chains is going to be per-

formed. In figure 4.7 the traces of the generated chains are represented for the population mean values. These seem to mix rapidly over the sample space of the posterior distributions. Therefore, all four parameters appear to have converged to their stationary distribution, so that sampling from the corresponding posterior distributions seems to be achieved.

Stationarity has been checked visually in the traces for each of the other variables in the model. All showed a well-mixing behaviour.

## 4.8 Conclusions

The tools introduced in previous chapters have proved to be remarkably appropriate to estimate concentration levels of cyclosporin in blood plasma following an infused multi-dose treatment. It has, therefore, been described a framework that links drug dosing with drug blood concentration levels over time with significant accuracy.

In what follows, drug concentration levels, and therefore particular dosing histories, will be linked to actual drug effect in order to gain further insight towards rational treatment.

# Chapter 5

# Cell Kinetics

A procedure that has been validated in the previous chapter as a flexible and simple framework to model the time course of drug concentration in blood plasma following drug administration has been introduced in chapter 3. In order to achieve the ultimate goal of this work (*i.e.* to relate drug dosing with changes in particular biological processes), the relationship between plasma drug concentration levels and changes in a chosen biological process needs to be studied.

In this work, the effects that drugs initiate in malignant cell populations will be analysed. In order to identify these effects correctly, the kinetics that govern the proliferation of cells in a drug-free context need to be studied first. This will be performed in this chapter.

Cell populations have historically been analysed in 'Continuous Time Markov Processes' theory (Bailey, 1967) and in 'Branching Processes' theory (Jagers, 1975). In this chapter both kinds of stochastic processes will be briefly introduced and corresponding models built.

Any modelling attempt should start by defining the physiological processes that it is intended to analyse. Section 5.1 constitutes a biological description

of the kinetics of individual cells, so that the foundations and assumptions of the models that follow will become apparent.

In section 5.2 a deterministic approach to model the cell cycle, and therefore the cell population, will be taken and discussed. A multi-type cell population model is proposed and the pros and cons expounded.

Section 5.3 starts by introducing the theory of 'Continuous Time Markov Processes'. 'Birth-and-Death' processes are explained and extended to accommodate different kinetic patterns within the population. Different types of cells are alowed within the population. The distributional constraints derived from vital assumptions required in this approach will be palliated to some 'weak' extent in the model proposed in section 5.4 by means of analysing a less ambitious kinetic pattern.

The theory of 'Branching Processes', discussed in section 5.5, constitutes an alternative approach to model cell populations. A population of cells of the same type are only considered. The implications of the required assumptions are physiologically less restrictive. Unfortunately, these processes are mathematically more complicated and rely on asymptotic results. An interesting 'easy-to-work-with' general approximation of one of the main limit results will be derived and illustrated. This simplifies considerably the use of the theory of 'Branching Processes' in the study of the evolution of cell populations.

## 5.1 Cell Cycle

The number of cells in a tumour or in the population of leukaemic cells in the circulatory system grows because the cell division rate is greater than that of death. They fail to respond effectively to the homeostatic control mechanisms which maintain the appropriate number of cells in normal renewal tissues (for

a detailed discussion of the cell kinetics of normal tissues see, for instance, Burgess and Nicola, 1983, Wright and Alison, 1984, Murakami *et al.*, 1985).

For decades it has been well known that there is in the life of a cell a well defined period during which DNA is synthesised. The concept of cell cycle was first introduced by Howard and Pelc (1953). More recently, experiments using different sophisticated techniques (Murray, 1993, Vincent, 1990) have led to a general model for the cell cycle. This model is represented in figure 5.1.

Within a cell population, not all the cells follow the kinetics of cell division. A new born cell enters in a phase called $G_1$ where the fate of the cell will be established. The $G_1$ period is usually longer than a third of the life length. During this early phase, three possible directions can be taken.

- The *natural* path to follow is to, after some time of maturation, enter in the $S$ phase, a very active phase where DNA duplication is undertaken. Two identical cells are *fertilized* during this period of considerable duration. In the subsequent stage, known as $G_2$ phase, the cell will get ready for the last and most dramatic stage of a cell's life, *mitosis* or $M$ phase. During *mitosis*, the cell will split into two newborn daughter cells, the reproducing cell ceasing to exist as such.

- The cell can enter a 'quiescent' or 'resting' state, known as the $G_0$ phase, at any time whilst in $G_1$. Cells will return to cycling at some indeterminate later time.

- The cell can also be terminally differentiated and cease proliferation. In this case, even though the death is imminent, the actual disappearance will not occur until the cycle is completed, so that the cell will go on with the cycle but without any kind of activity. We call this, the $I$ period to show that the cells are irreversibly out of the cycle.

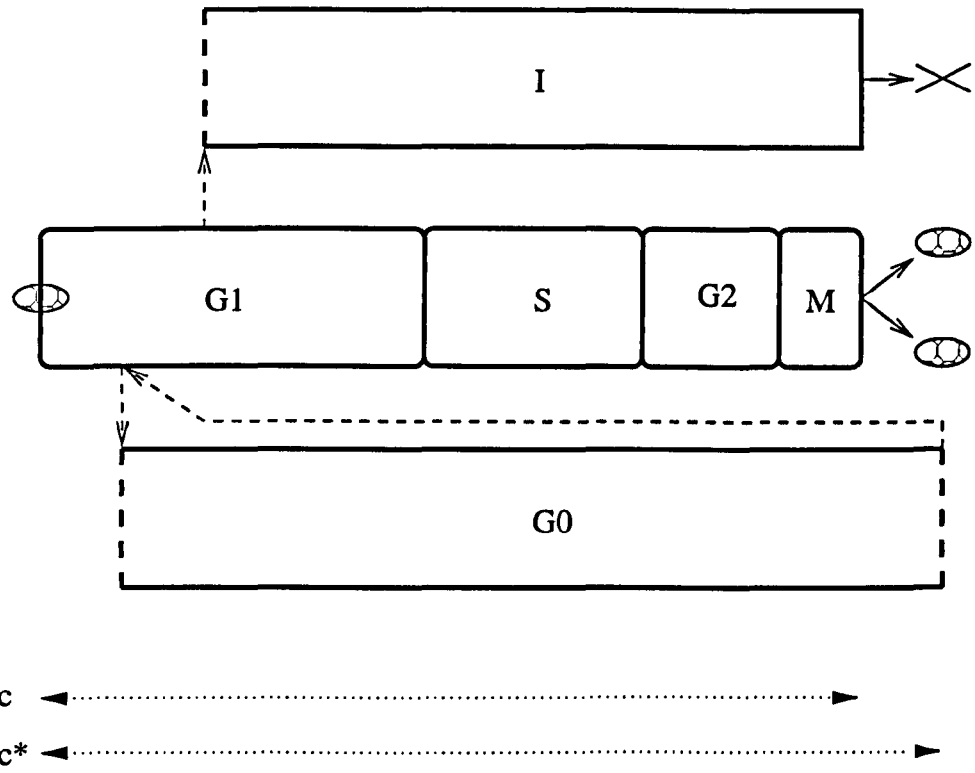



Figure 5.1: Cell Cycle. A dashed vertical edge or arrow means that the cell has got the option to take that path. Solid vertical edges or arrows, instead, indicate an unavoidable path. $T_c$ is the total cycle time if the cell is proliferative (i.e., $G_1 \rightarrow S \rightarrow G_2 \rightarrow M$) or is irreversibly differentiated (i.e., $G_1 \rightarrow I$). If the cell enters a resting phase (i.e., $G_1 \rightarrow G_0 \rightarrow S \rightarrow G \rightarrow M$), then the cell cycle time will be $T_{c^*}$.

Cells that *choose* to take the first path are called *proliferative*, whereas cells that enter into a resting phase are known as *non proliferative* or *resting* cells. Finally, we will say that the cells that are committed to die and leave the cycle are *irreversibly differentiated.*

If we denote the times spent in each phase by $T_{G_1}$, $T_S$, $T_{G_2}$, $T_M$, $T_I$ and $T_{G_0}$ respectively, the total cycle time $T_c$ will be given by:

$$T_c = \begin{cases} T_{G_1} + T_S + T_{G_2} + T_M & \text{if proliferative or goes to } I \\ T_{G_1} + T_{G_0} + T_S + T_{G_2} + T_M & \text{if goes to } G_0 \end{cases} . \quad (5.1)$$

Each of these quantities is not fixed and varies substantially between cells. Therefore, it is natural to treat them as random variables. The variable which is known to be the most unpredictable and obscure one is $T_{G_0}$. In some models, $T_{G_0}$ has been considered to be infinity (Steel, 1968, Steel, 1977). The consequences of noncycling cells in modelling and analysis of experiments are profound (White and Terry, 2000).

## 5.2 Deterministic Approach to model the Cell Cycle

### 5.2.1 Introduction

We are concerned with a population of cells following the behaviour described in the previous section. Interest lies on understanding both qualitatively and quantitatively the growth of such populations.

A high variability has been acknowledged in most of the mathematical descriptions of biological populations. The work by Elton (1958) was followed by a vast literature on modelling the spread of biological populations. For a review

of different models applied in different branches of biology see , for example, Hengeveld (1989), Renshaw (1991) and references therein.

The deterministic approach has proved validity in the study and prediction of certain large populations where the variability is regarded as relatively small enough. Bailey (1967, p. 2) states that there is nothing inappropriate in predicting that the present population of, say, 51226 persons in a town will have grown to 60863 in ten years' time, provided the latter estimate is correct to within a few hundred units. But if, in a small family of four children, one of whom develops measles today, we predict the total number of cases in a week's time as 2.37, it is most unlikely that this figure will have any particular significance. The prediction of small populations like this should, therefore, be based on a probabilistic approach.

Starting with one single cell, the number of cells in a population after certain time period will vary substantially between different populations. Nevertheless, the first attempt in modelling the fate of large populations has historically been to ignore such variation and fix terms such as 'birth rate' and 'death rate' as continuous and steady flows. Given these constant rates, as it will be shown in what follows, the population size can be determined at any time as the solution of the corresponding set of differential equations. This constitutes the deterministic approach.

The simplest deterministic model to predict the number of individuals in a population with one ancestor and with a steady net growth rate of $\lambda$ would be the solution of the differential equation

$$\frac{dX(t)}{dt} = \lambda X(t),$$

where X(t) is the positive integer-valued variable which counts the number of individuals alive at time $t \geq 0$. Therefore, the number of individuals at time $t$ will be given by $X(t) = e^{\lambda t}$.

### 5.2.2 Cell Populations

As introduced in the previous section, there are three types of cells depending on the role they undertake within the cell cycle. In figure 5.2 a population consisting of these three different subpopulations with the corresponding deterministic migration rates is drawn.

Denoting by the positive integer-valued variables $X_p(t)$, $X_r(t)$ and $X_i(t)$ the number of proliferative, resting and irreversibly out of cycle cells at time $t \geq 0$ respectively, and by $k_{pp}$, $k_{pr}$, $k_{rp}$, $k_{pi}$ and $k_d$ the migration rates shown in figure 5.2, the dynamics of such a population would be given by the following set of differential equations:

$$\frac{dX_p(t)}{dt} = (k_{pp} - k_{pr} - k_{pi})X_p(t) + k_{rp}X_r(t) \tag{5.2}$$

$$\frac{dX_r(t)}{dt} = k_{pr}X_p(t) - k_{rp}X_r(t) \tag{5.3}$$

$$\frac{dX_i(t)}{dt} = k_{pi}X_p(t) - k_dX_i(t) \tag{5.4}$$

The solution set of this system specifies exactly the number of cells of each type at any instant of time, and therefore, the total number of cells in the population:

$$X_p(t) = C_p(e^{\lambda_1 t} + e^{\lambda_2 t}) \tag{5.5}$$

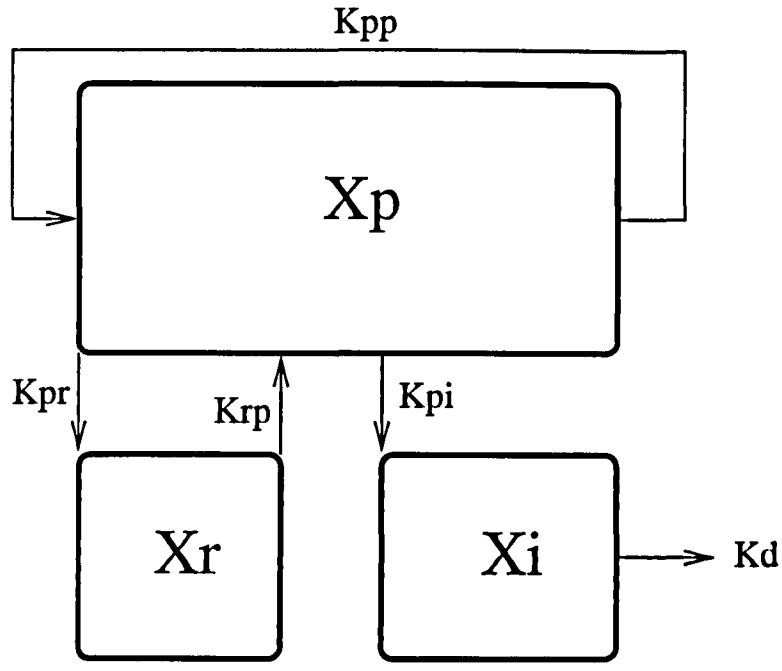
$$X_r(t) = C_r(-\frac{\alpha - \lambda_1}{k_{rp}}e^{\lambda_1 t} - \frac{\alpha - \lambda_2}{k_{rp}}e^{\lambda_2 t}) \tag{5.6}$$

$$X_i(t) = C_i(\frac{k_{pi}}{k_d + \lambda_1}e^{\lambda_1 t} + \frac{k_{pi}}{k_d + \lambda_2}e^{\lambda_2 t} + e^{-k_d t}) \tag{5.7}$$

where $\alpha = k_{pp} - k_{pr} - k_{pi}$, and defining $D = (\alpha - k_{rp})^2 + 4k_{rp}(k_{pr} + \alpha) > 0$, $\lambda_{1,2} = \frac{\alpha - k_{rp} \pm \sqrt{D}}{2}$ respectively. The constants $C_p$, $C_r$ and $C_i$ are defined by the initial values. When analysing population with a single ancestor, these initial values are given by: $X_p(0) = 1$, $X_r(0) = X_i(0) = 0$.

A deterministic model like the one presented, as discussed above, could be of legitimate use in the case of large populations with small variability as an

Figure 5.2: Multi-Type cell population. There are three different types of cells, the proliferative cells, $X_p$, the resting cells, $X_r$, and the irreversibly differentiated cells, $X_i$, with migration rates $k_{pr}$,$k_{rp}$,$k_{pi}$ and $k_d$ and splitting rate $k_{pp}$.




approximate trend. If we start to model a large population of cells where a count of cells of each type is available and there is a good knowledge of the migration rates (or are estimated correctly), we could predict fairly safely the evolution of the population with equations (5.5)-(5.7) up to a time point where the assumed variability becomes considerable in relation to the number of cells. At this time more realistic stochastic models should be used instead, as well as if we were analysing the early stages of the population.

It should be kept in mind that stochastic models are not merely generalisations of their deterministic counterparts. The properties of each of the models can vary substantially. In many cases the means of the stochastic models will turn out to be equal to the corresponding exact value of the deterministic growth models. But, this is by no means always true.

## 5.3 Continuous Time Markov Processes

### 5.3.1 Introduction

We aim to build a statistical model for the growth of cell populations. We start by defining the random variable $X(t)$, taking nonnegative integer values, representing the size of the population at time $t$. The time parameter will be continuous as changes in the population could happen at any instant.

There are different stochastic processes that could be applied to study the properties and evolution of $\{X(t), t \geq 0\}$. One of the simplest but still the most broadly used are the well known Markov Processes in Continuous Time. Many authors have used Markov processes as mathematical models for the growth of biological or physical populations (see, for instance, for general references Bailey, 1967, Chiang, 1980 and Ross, 1996 and references therein; for particular applications, for instance, Lefevre, 1981, Metz *et al.*, 1983, Norberg, 1995, Chick *et al.*, 1996, Jaiswal *et al.*, 1997).

Although too restrictive and simplistic in many cases, they are of great interest as they can lead to analytic results that can give important insights about the real process underneath.

We say that the stochastic process $\{X(t), t \geq 0\}$ is a *continuous time Markov chain* if for all $s, t \geq 0$, and nonnegative integers $i, j, x(u), 0 \leq u \leq s$,

$$\begin{aligned} P\{X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u < s\} \\ = P\{X(t+s) = j | X(s) = i\}. \end{aligned} \tag{5.8}$$

In other words, if the size of the population is known at any time, say $X(t) = i$, then given $X(t) = i$, the subsequent behaviour of the population is assumed to be independent of the past history. This is the *Markov property.*

If, in addition, $P\{X(t+s) = j | X(s) = i\}$ is independent of $s$, then the

continuous-time Markov chain is said to have stationary or homogeneous transition probabilities.

We aim to find out the probability distribution function of the process $\{X(t), t \geq 0\}$ or, at least, to be able to capture some information from it (e.g., the expected value). Let $X(t)$ be an positive integer-valued random variable. If we denote by $p_n(t)$ the probability that $X(t)$ takes the value $n$ at time $t$, the probability-generating function, the moment-generating function and the cumulant-generating function are respectively given by:

$$P(x,t) = E(x^{X(t)}) \tag{5.9}$$

$$M(\theta,t) = E(e^{\theta X(t)}) = 1 + \sum_{r=1}^{\infty} \mu'_r(t)\theta^r/r! \tag{5.10}$$

$$K(\theta,t) = \ln(M(\theta,t) = \sum_{r=1}^{\infty} \kappa_r(t)\theta^r/r! \tag{5.11}$$

where $\mu'_r(t)$ is the $r^{th}$ moment about the origin and $\kappa_r(t)$ the $r^{th}$ cumulant.

We will follow what Bailey (1967, pp. 70 – 74) terms the 'random-variable technique' for writing down, in an intuitive way, the corresponding partial differential equations for the probability-generating function or the moment-generating function, and therefore, for the cumulant-generating function of the process.

In some applications these partial differential equations will be straightforward to solve, whereas in most of the practical applications this will be a complicated or even an impossible task. Nevertheless, as we will see below, even though in most cases explicit solutions for the distribution at time $t$ can not be found, it is relatively simple to find the corresponding moments or cumulants.

One general method of approximating distributions is first to solve for the cumulants of the process and then to find suitable approximating distributions

with matching cumulants (see, for instance, Renshaw, 1991, Wehrly *et al.*, 1993 or Renshaw, 2000).

Let the change in the population $X(t)$ during the interval $\Delta t$ be defined as

$$\Delta X(t) = X(t + \Delta t) - X(t). \tag{5.12}$$

Usually there is a finite number of possible transitions during $\Delta t$. Hence, let us suppose that we can write

$$P\{\Delta X(t) = j | X(t)\} = f_j(X)\Delta t, \quad j \neq 0, \tag{5.13}$$

where $f_j(X)$ are appropriate nonnegative functions of $X(t)$ and $j$ may be a positive or negative integer (in most applications $f_j$ will be very simple function). The probability of no change in $X(t)$ during $\Delta t$ is therefore given by

$$P\{\Delta X(t) = 0 | X(t)\} = 1 - \sum_{j \neq 0} f_j(X)\Delta t. \tag{5.14}$$

Applying expectations in equation (5.12), it is easy to show that for an arbitrary function $\Phi\{X(t)\}$, we can write

$$E_{t+\Delta t}[\Phi\{X(t + \Delta t)\}] = E_t\{E_{\Delta t|t}[\Phi\{X(t) + \Delta X(t)\}]\}, \tag{5.15}$$

where $E_t$ and $E_{t+\Delta t}$ are the expectations of the function $\Phi$ at time $t$ with respect to variations in $X(t)$, and time $t + \Delta t$ for variations in $X(t + \Delta t)$ respectively. And, $E_{\Delta t|t}$ is the conditional expectation at the end of the interval $\Delta t$ given the value of $X(t)$ at time $t$.

And, in particular, for $M = \Phi$, where $M(\theta, t)$ is the moment generating function we get:

$$\begin{aligned} M(\theta, t + \Delta t) &= E_{t+\Delta t}\{e^{\theta X(t+\Delta t)}\} && (5.16) \\ &= E_{t+\Delta t}\{e^{\theta X(t)+\theta \Delta X(t)}\} && (5.17) \\ &= E_t[e^{\theta X(t)} E_{\Delta t|t}\{e^{\theta \Delta X(t)}\}]. && (5.18) \end{aligned}$$

Differentiating $M(\theta,t)$ with respect to $t$, and using equation (5.18), it follows that,

$$\frac{\partial M(\theta,t)}{\partial t} = \lim_{\Delta t\to 0}\frac{M(\theta,t+\Delta t)-M(\theta,t)}{\Delta t} \tag{5.19}$$

$$= \lim_{\Delta t\to 0}\frac{1}{\Delta t}[E_t\{e^{\theta X(t)}E_{\Delta t|t}\{e^{\theta\Delta X(t)}\}\}-E_t\{e^{\theta X(t)}\}] \tag{5.20}$$

$$= E_t[e^{\theta X(t)}\lim_{\Delta t\to 0}E_{\Delta t|t}\{\frac{(e^{\theta\Delta X(t)}-1)}{\Delta t}\}]. \tag{5.21}$$

If the limit as $\Delta t\to 0$ of the expectation inside the brackets in equation (5.21) is $\Psi(\theta,t,X)$, then this equation can be written as:

$$\frac{\partial M(\theta,t)}{\partial t} = E_t\{e^{\theta X(t)}\Psi(\theta,t,X)\} \tag{5.22}$$

$$= \Psi(\theta,t,\frac{\partial}{\partial\theta})M(\theta,t), \tag{5.23}$$

where the operator $\partial/\partial\theta$ acts only on $M(\theta,t)$.

In the special case in which the conditional transition probabilities are given by equations (5.13) and (5.14), the function $\Psi$ has the form

$$\Psi(\theta,t,X) = \lim_{\Delta t\to 0}\frac{\{1-\sum_{j\neq 0}f_j(X)\Delta t\}+\sum_{j\neq 0}f_j(X)\Delta te^{j\theta}-1}{\Delta t} \tag{5.24}$$

$$= \sum_{j\neq 0}(e^{j\theta}-1)f_j(X). \tag{5.25}$$

Using equation (5.25), the differential equation (5.23) simplifies to

$$\frac{\partial M(\theta,t)}{\partial t} = \sum_{j\neq 0}(e^{j\theta}-1)f_j(X)\frac{\partial M(\theta,t)}{\partial\theta}. \tag{5.26}$$

And, therefore, it follows that,

$$\frac{\partial K(\theta,t)}{\partial t} = \sum_{j\neq 0}(e^{j\theta}-1)f_j(X)\frac{\partial K(\theta,t)}{\partial\theta}, \tag{5.27}$$

$$\frac{\partial P(x,t)}{\partial t} = \sum_{j\neq 0}(x^j-1)f_j(X)x\frac{\partial P(x,t)}{\partial x}. \tag{5.28}$$

These differential equations can be solved analytically for some stochastic processes for which exact analytical solution for the probability distribution at time $t$ can not be found.

In what follows, for all the processes we will present, the corresponding differential equations (5.26), (5.27) or (5.28) will be calculated in order to capture the moments of the underlying distributions.

### 5.3.2 Linear Birth and Death Processes

Let us start with the most familiar time continuous Markov process, the 'Linear Birth and Death Process'. First introduced by Yule (1924), this models have been used in a diverse variety of scenarios (see, for instance, Harris, 1963, Bailey, 1967, Athreya and Ney, 1972, Jagers, 1975, and Chiang, 1980, for different applications and further references).

At time $t$ there is a population of cells whose total number is given by the discrete variable $X_p(t)$ [1], where the probability that $X_p(t)$ takes the positive integer value $n$ is given by $p_n(t)$. Each cell lives for certain time and at its death it can be either replaced by two new born daughter cells or by none. In the former case we will say there has been a birth (twin birth, the reproducing cell dying) and in the latter one that there has been a death. If there is a birth, the total population will increase by one unit, whereas if there is death, it will decrease by one unit.

Let us start off with the assumption that the probability of a given cell dividing in time $\Delta t$ is $\lambda \Delta t$, and that the probability of dying in $\Delta t$ is given by $\mu \Delta t$. This is equivalent to assuming that the life length $u$ of the cell has an exponential distribution with parameter $(\lambda + \mu)$ so that

$$l(u) = (\lambda + \mu) e^{-(\lambda+\mu)u}, \quad 0 \leq u < \infty. \tag{5.29}$$

It is convenient to keep in mind the importance of the property of *lack of memory* of the Exponential Distribution. This means that the probability of

[1] Subscript $p$ refers to proliferative cells.

an event occurring in the immediate future is independent of the time of the last occurrence. In terms of cell life lengths, this property implies that the probability that a cell will be alive at time $s+l \geq 0$, given that it has survived $s \geq 0$, is the same as the initial probability that it lives at $l \geq 0$. In other words, if a cell is alive at time $s$, then the distribution of its remaining life is the original lifetime distribution. Therefore, the lifetime of a cell is considered to be completely random. This assumption will be relaxed in following sections.

The population will be composed of indistinguishable cells. All the events that can happen to a cell during the interval $(t, t+\Delta t)$ are independent of the event happening to other cells and of the events that happened to the cell in the past.

As we assume that the lifetimes of the cells are independent, exponentially distributed random variables, the possible changes in the population size $X_p(t) = n$ at time $t$ can be classified as follows:

- Prob. of one birth $= \lambda n \Delta t + o(\Delta t)$
- Prob. of one death $= \mu n \Delta t + o(\Delta t)$
- Prob. of more than one of these events $= o(\Delta t)$
- Prob. of no jump $= 1 - (\lambda n + \mu n)\Delta t + o(\Delta t)$.

Note that in this process the probability of a jump (either a birth or a death) during a short time interval $\Delta t$ depends on the present population size but not on any other function of the time. Thus the process is homogeneous with respect to time.

Given that $X_p(t) = n$, the waiting time $w$ between jumps has the exponential distribution

$$q(w) = (\lambda + \mu) n e^{-(\lambda+\mu)nw} \quad 0 \leq w < \infty. \tag{5.30}$$

If a jump actually occurs, the probability of a birth is $\lambda/(\lambda+\mu)$ (population increases in one unit) and that of a death $\mu/(\lambda+\mu)$ (population decreases in one unit).

For this first simple model, the probability distribution function is straightforward and we do not need to use the 'random-variable technique'. It follows immediately that,

$$\begin{aligned} p_n(t+\Delta t) &= p_n(t)(1-\lambda n\Delta t-\mu n\Delta t)+p_{n-1}(t)(\lambda(n-1)\Delta t) \\ &\quad +p_{n+1}(t)(\mu n\Delta t)+o(\Delta t). \end{aligned} \tag{5.31}$$

It is easy to check (see, for instance, Bailey, 1967, pp. 92–97) that, if we start off with $X_p(0)=a$, the corresponding probability distribution is given by

$$p_n(t)=\begin{cases} \sum_{j=0}^{min(a,n)} \binom{a}{j}\binom{a+n-j-1}{a-1}\alpha^{a-j}\beta^{n-j}(1-\alpha-\beta)^j & \text{if } n\geq 1 \\ \alpha^a & \text{if } n=0 \end{cases}, \tag{5.32}$$

where

$$\alpha=\frac{\mu(e^{(\lambda-\mu)t}-1)}{\lambda e^{(\lambda-\mu)t}-\mu}, \quad \beta=\frac{\lambda(e^{(\lambda-\mu)t}-1)}{\lambda e^{(\lambda-\mu)t}-\mu}. \tag{5.33}$$

In the special case where $a=1$, this simplifies to:

$$p_n(t)=\begin{cases} (1-\alpha)(1-\beta)\beta^{n-1} & \text{if } n>0 \\ \alpha & \text{if } n=0 \end{cases}, \tag{5.34}$$

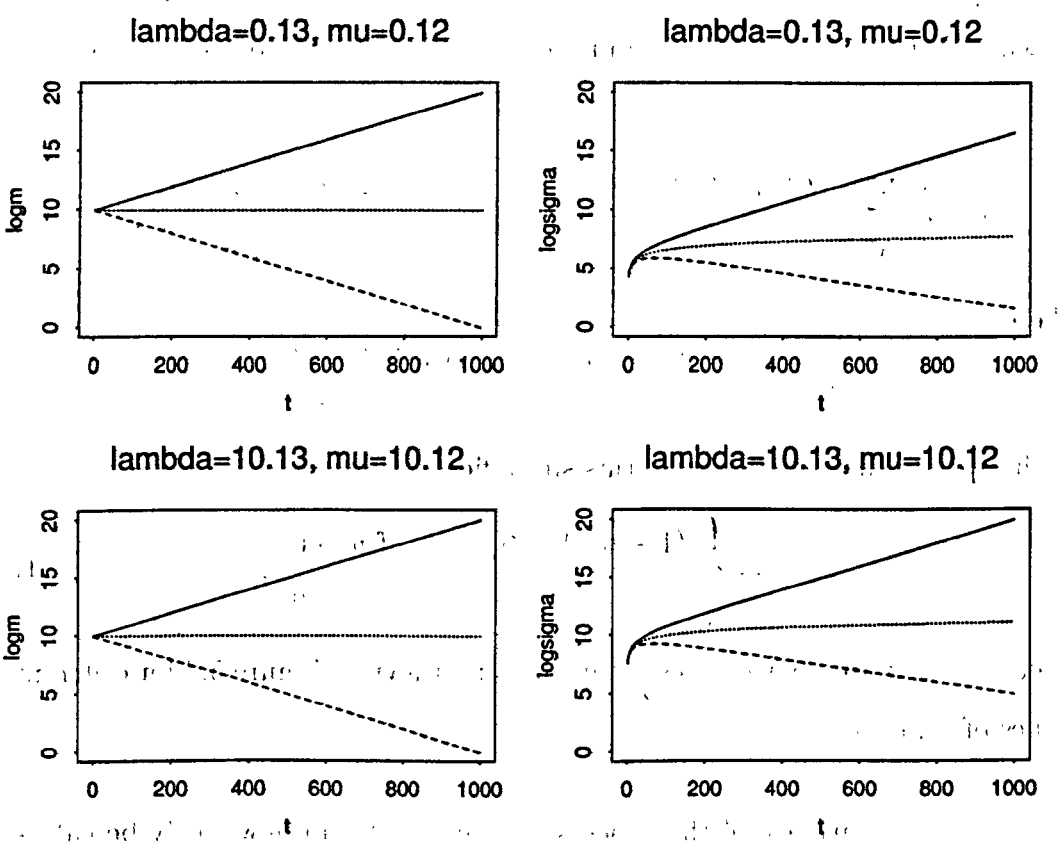
where $\alpha$ and $\beta$ are above. These moments are drawn in Figure 5.3 for different values of $\lambda$ and $\mu$.

The mean and variance of the process for any $a>0$ can now easily be calculated as:

$$m(t)=\begin{cases} ae^{(\lambda-\mu)t} & \text{if } \lambda\neq\mu \\ a & \text{if } \lambda=\mu \end{cases}, \tag{5.35}$$

$$\sigma^2(t)=\begin{cases} \frac{a(\lambda+\mu)}{(\lambda-\mu)}e^{(\lambda-\mu)t}(e^{(\lambda-\mu)t}-1) & \text{if } \lambda\neq\mu \\ 2a\lambda t & \text{if } \lambda=\mu \end{cases}. \tag{5.36}$$

Figure 5.3: Mean (left plots) and Standard Deviation (right plots) in logarithms for the Birth and Death Process. In the superior plots, straight lines correspond to $\lambda = 0.13$ and $\mu = 0.12$, dashed lines to $\lambda = 0.125$ and $\mu = 0.125$, and dotted lines to $\lambda = 0.12$ and $\mu = 0.13$. In the inferior plots these values are 10.13 and 10.12, 10.125 and 10.125, and 10.12 and 10.13 respectively.



It should be pointed out that the expected value is nothing but the growth rate of the corresponding deterministic model and that it will only depend on the difference between the instantaneous probabilities $\lambda - \mu$ and the initial value.

In figure 5.3 the expected value and the standard deviation is compared for two different life length distributions under three different birth and death probability scenarios in a population that starts with 20000 cells.

In the superior plots, the life length of each of the cells is Exponentially distributed with parameter $\lambda+\mu = 0.25$, so that the mean value and the standard deviation of the cell length are given by 4. After a cell dies, we have considered three different birth and death probabilities:

- $\lambda = 0.13$ and $\mu = 0.12$, so that with probability 0.52 there will be one more cell in the population.
- $\lambda = 0.125$ and $\mu = 0.125$, so that the population will be unchanged.
- $\lambda = 0.12$ and $\mu = 0.13$, so that with probability 0.52 there will be one cell less in the population.

In the inferior plots, the life length of each of the cells is Exponentially distributed with parameter $\lambda+\mu = 10.13+10.12 = 20.25$, so that the mean value and deviation of the cell length are given by 0.049383. In this case, the three probability scenarios are given by:

- $\lambda = 10.13$ and $\mu = 10.12$, so that with probability 0.50025 there will be one more cell in the population.
- $\lambda = 10.125$ and $\mu = 10.125$, so that the population will be unchanged.

- $\lambda = 10.12$ and $\mu = 10.13$, so that with probability 0.49975 there will be one cell less in the population.

Figure 5.3 confirms that the expected value of the population only depends on the difference $\lambda - \mu$.

In this very simple model, where the probability distribution of the stochastic process is easily attained, it is straightforward (see, for instance, Bailey, 1967, pp.95 – 96)to calculate the extinction probability of the population:

$$\lim_{t\to\infty} p_0(t) = \begin{cases} 1 & \text{if } \lambda \leq \mu \\ (\frac{\mu}{\lambda})^a & \text{if } \lambda > \mu \end{cases} . \tag{5.37}$$

Therefore, the chance of extinction of the population is positive for any value of $\lambda$ and $\mu$. Exctinction will actually be certain unless the birth rate exceeds the death rate, in which case it will happen with probability $(\frac{\mu}{\lambda})^a$.

### 5.3.3 Multi-Type Birth and Death Processes

In the previous section, cells within the population were indistinguishable as they all followed independently the same probabilistic behaviour with exponential life times. Several extensions of this process are possible.

The first natural generalisation of the birth-and-death process is to allow different probabilistic behaviours for different distinguishable subpopulations. These processes are called Multi-Type Brith and Death Processes provided that each subpopulation has an exponential life length distribution (see, for instance Bailey, 1967, and references therein).

In this section we are going to propose a new model that will enable us to follow the evolution of different cell types. Therefore, two processes which remain hidden in simple Birth and Death processes will be taken into account:

cells that reversibly enter into a resting phase and cells that are irreversibly differentiated but still remain for certain time in the organism. These two processes, when relevant, will be of substantial importance when analysing the whole population.

As we have already seen previously in this chapter, within a cell population there are three types of different cells depending on the cycle kinetics they follow, i.e. proliferative, resting and irreversibly out of the cycle. It seems therefore, necessary to assign different probabilistic behaviours to each of these cell groups.

At time $t$, there are three subpopulations of different cell types. The advantage of this model is that at any time $t$ we will be able to count the number of cells in each of these subpopulations that follow different kinetics. Therefore, we will be able not only to get better insight of the evolution of the whole population but also to introduce changes in the kinetics of a particular subpopulation (for instance, by increasing the ratio of proliferative cells that get differentiated irreversibly introducing drugs). The deterministic counterpart of the model we propose has been presented in figure 5.2.

There is a pool of proliferative cells, $X_p(t)$. Given the migration and splitting rates shown in Figure 5.2, during the short period $\Delta t$, each proliferative cell can move towards mitosis, and therefore split into two new born daughter cells with probability $k_{pp}\Delta t$, can decide the leave the cycle reversibly, *i.e.* go resting, with probability $k_{pr}\Delta t$ or can leave the cycle irreversibly so that it will eventually disappear, with probability $k_{pi}\Delta t$.

The number of resting cells at time $t$ is given by the random variable $X_r(t)$. Within the interval $\Delta t$ a resting cell will become proliferative again with probability $k_{rp}\Delta t$.

Finally, $X_i(t)$ is the total number of cells which are irreversibly differentiated. Cells in this subpopulation, will be expelled from the organism during $\Delta t$ with probability $k_d \Delta t$.

All the events that can happen to a cell within a subpopulation during the interval $(t, t + \Delta t)$ are assumed to be independent of the events happening to other cells of the same subpopulations and of the event that happened to the cell in the past. This implies that life lengths within a subpopulation are independent and identically distributed exponential random variables.

From the assumptions above, it follows that there are five possible types of changes over a short time period $\Delta t$ in the subpopulation sizes. The *instantaneous* probabilities that these changes occur in the interval $(t, \Delta t)$ are consequently given by:

- Prob. $X_p$ increases by one unit (cell division) $= k_{pp} X_p(t) \Delta t + o(\Delta t)$.
- Prob. $X_p$ decreases by one unit and $X_r$ increases by one unit (a proliferative cell goes resting) $= k_{pr} X_p(t) \Delta t + o(\Delta t)$.
- Prob. $X_p$ increases by one unit and $X_r$ decreases by one unit (a resting cell returns to the cycle) $= k_{rp} X_r(t) \Delta t + o(\Delta t)$.
- Prob. $X_p$ decreases by one unit and $X_i$ increases by one unit (cell differentiation) $= k_{pi} X_p(t) \Delta t + o(\Delta t)$.
- Prob. $X_i$ decreases by one unit (cell death) $= k_d X_i(t) \Delta t + o(\Delta t)$.

If the joint probability distribution of $X_p(t)$, $X_r(t)$ and $X_i(t)$ at time $t$ is given by:

$$P\{X_p(t) = n_p, X_r(t) = n_r, X_i(t) = n_i\} = p_{n_p n_r n_i}(t), \tag{5.38}$$

the probability generating function, the moment generating function and the cumulant generating function are, respectively, given by:

$$P(x_p, x_r, x_i) = \sum_{n_p,n_r,n_i} p_{n_p,n_r,n_i}(t) x_p^{n_p} x_r^{n_r} x_i^{n_i}, \tag{5.39}$$

$$M(\theta_p, \theta_r, \theta_i, t) = P(e^{\theta_p}, e^{\theta_r}, e^{\theta_i}) \tag{5.40}$$

$$= \sum_{n_p,n_r,n_i} p_{n_p,n_r,n_i}(t) e^{n_p\theta_p + n_r\theta_r + n_i\theta_i} \tag{5.41}$$

$$K(\theta_p, \theta_r, \theta_i, t) = logM(\theta_p, \theta_r, \theta_i, t) \tag{5.42}$$

Expanding $K(\theta_p, \theta_r, \theta_i, t)$ in powers of $\theta_p, \theta_r$ and $\theta_i$ we get it in terms of joint moments or cumulants, i.e.

$$K(\theta_p, \theta_r, \theta_i, t) = \sum_{u_p,u_r,u_i} \kappa_{u_p u_r u_i}(t) \frac{\theta_p^{u_p} \theta_r^{u_r} \theta_i^{u_i}}{u_p! u_r! u_i!}, \tag{5.43}$$

where $k_{000} \equiv 0$.

Now, let us suppose that the joint probability distribution of relevant transitions in the interval $\Delta t$ is given by:

$$\begin{aligned} P\{\Delta X_p(t) = j_p, \Delta X_r(t) = j_r, \Delta X_i(t) = j_i \mid X_p(t), X_r(t), X_i(t)\} \\ = f_{j_p j_r j_i}(x_p, x_r, x_i)\Delta t, \end{aligned} \tag{5.44}$$

where $j_p, j_r, j_i$ are not all zero and that the only possible transitions are:

- $f_{100} = k_{pp} X_p(t)$
- $f_{1-10} = k_{rp} X_r(t)$
- $f_{-110} = k_{pr} X_p(t)$
- $f_{-101} = k_{pi} X_p(t)$
- $f_{00-1} = k_d X_i(t)$

Extending the results presented on section 5.3.1, to the present multi- dimensional case, it is easy to see that the differential equation for the cumulant-generating function given in equation (5.27) can be now written as follows.

$$\frac{\partial K}{\partial t} = ([e^{\theta_p} - 1]k_{pp} + ([e^{-\theta_p + \theta_r} - 1]k_{pr} + [e^{-\theta_p + \theta_i} - 1]k_{pi}) \frac{\partial K}{\partial \theta_p}$$

$$+ [e^{\theta_p - \theta_r} - 1]k_{rp}\frac{\partial K}{\partial \theta_r} + [e^{-\theta_i} - 1]k_d\frac{\partial K}{\partial \theta_i} \tag{5.45}$$

Differentiating equation (5.43) with respect to $t$:

$$\frac{\partial K}{\partial t} = \sum_{u_p,u_r,u_i} \frac{d\kappa_{u_p u_r u_i}}{dt}\frac{\theta_p^{u_p}\theta_r^{u_r}\theta_i^{u_i}}{u_p!u_r!u_i!}, \tag{5.46}$$

Equating coefficients of $\theta_p$, $\theta_r$ and $\theta_i$ on both sides of equations (5.45) and (5.46) we get the following set of differential equations for the expected values of the process:

$$\frac{d\kappa_{100}}{dt} = (k_{pp} - k_{pr} - k_{pi})\kappa_{100} + k_{rp}\kappa_{010} \tag{5.47}$$

$$\frac{d\kappa_{010}}{dt} = k_{pr}\kappa_{100} - k_{rp}\kappa_{010} \tag{5.48}$$

$$\frac{d\kappa_{001}}{dt} = k_{pi}\kappa_{100} - k_d\kappa_{001}. \tag{5.49}$$

Therefore, the expected values of the number of cells of each type in this process will coincide with the corresponding deterministic counterparts presented in equations (5.5)-(5.7):

$$\kappa_p(t) = C_p(e^{\lambda_1 t} + e^{\lambda_2 t}) \tag{5.50}$$

$$\kappa_r(t) = C_r(-\frac{\alpha - \lambda_1}{k_{rp}}e^{\lambda_1 t} - \frac{\alpha - \lambda_2}{k_{rp}}e^{\lambda_2 t}) \tag{5.51}$$

$$\kappa_i(t) = C_i(\frac{k_{pi}}{k_d + \lambda_1}e^{\lambda_1 t} + \frac{k_{pi}}{k_d + \lambda_2}e^{\lambda_2 t} + e^{-k_d t}) \tag{5.52}$$

where the constants $\alpha$, $\lambda_1$, $\lambda_2$, $C_p$, $C_r$ and $C_i$ have been already defined for the deterministic model.

In a similar manner, equating coefficients of $\theta_p^2$, $\theta_r^2$, $\theta_i^2$, $\theta_p\theta_r$, $\theta_p\theta_i$ and $\theta_r\theta_i$ in equations (5.45) and (5.46), second order moments of the process can be easily

obtained as the solution of the following set of differential equations:

$$\begin{aligned}\frac{d\kappa_{200}}{dt} = {} & (k_{pp}+k_{pr}+k_{pi})\kappa_{100}+k_{rp}\kappa_{010}+2(k_{pp}-k_{pr}-k_{pi})\kappa_{200} \\ & +2\kappa_{rp}\kappa_{110},\end{aligned} \tag{5.53}$$

$$\frac{d\kappa_{020}}{dt} = k_{pr}\kappa_{100}+k_{rp}\kappa_{010}+2k_{pr}\kappa_{110}-2k_{rp}\kappa_{020}, \tag{5.54}$$

$$\frac{d\kappa_{002}}{dt} = k_{pi}\kappa_{100}+2k_{pi}\kappa_{101}+k_{d}\kappa_{001}-2k_{d}\kappa_{002}, \tag{5.55}$$

$$\begin{aligned}\frac{d\kappa_{110}}{dt} = {} & -k_{pr}\kappa_{100}-k_{rp}\kappa_{010}+(k_{pp}-k_{pi}-k_{pr})\kappa_{110}+k_{pr}\kappa_{200} \\ & +k_{rp}(\kappa_{020}-\kappa_{010}),\end{aligned} \tag{5.56}$$

$$\begin{aligned}\frac{d\kappa_{101}}{dt} = {} & -k_{pi}\kappa_{100}+(k_{pp}-k_{pi}-k_{pr})\kappa_{101}+k_{pi}\kappa_{200}+k_{rp}\kappa_{011} \\ & -k_{d}\kappa_{101},\end{aligned} \tag{5.57}$$

$$\frac{d\kappa_{011}}{dt} = k_{pi}\kappa_{110}. \tag{5.58}$$

In figure 5.4, the expected values of the number of cells of each type for different multi-type birth and death processes are represented. In the top-left plot proliferative cells can only either die or split. As the instantaneous probability of dividing ($k_{pp} = 0.06$) is greater than the probability of irreversible differentiation ($k_{pi} = 0.03$) with a short delay prior to actually dying ($k_d = 0.1$), the population grows rapidly. In this first case, the population is basically constituted by proliferative cells.

In the remaining plots, fixing $k_{pi}$ and $k_d$ as before, cells are allowed to enter into a resting phase. In all cases $k_{pr} = 0.01$ but with different resting times. In the top-right plot, cells rest for a relatively short time ($k_{rp} = 0.5$), so that the proportion of resting cells is very small. However, note that even with short resting periods, the whole population grows considerably slower than in the previous case. Longer resting times are considered in the remaining plots (actually, the resting time is infinity in the last case) and, as one would expect, the population growth decreases as the resting subpopulation increases due to this pool of non-proliferative cells.

Figure 5.4: Multi-Type Birth and Death Processes. Four different processes with $k_{pp}$, $k_{pr}$ and $k_{rp}$ as indicated. In all cases, $k_{pi} = 0.03$ and $k_d = 0.1$. Cumulative plots of the expected number of cells of each type in logarithms over time, starting with one proliferative cell. Straight lines correspond to proliferative cells, dotted lines to resting cells and dashed lines to irreversibly differentiated cells.

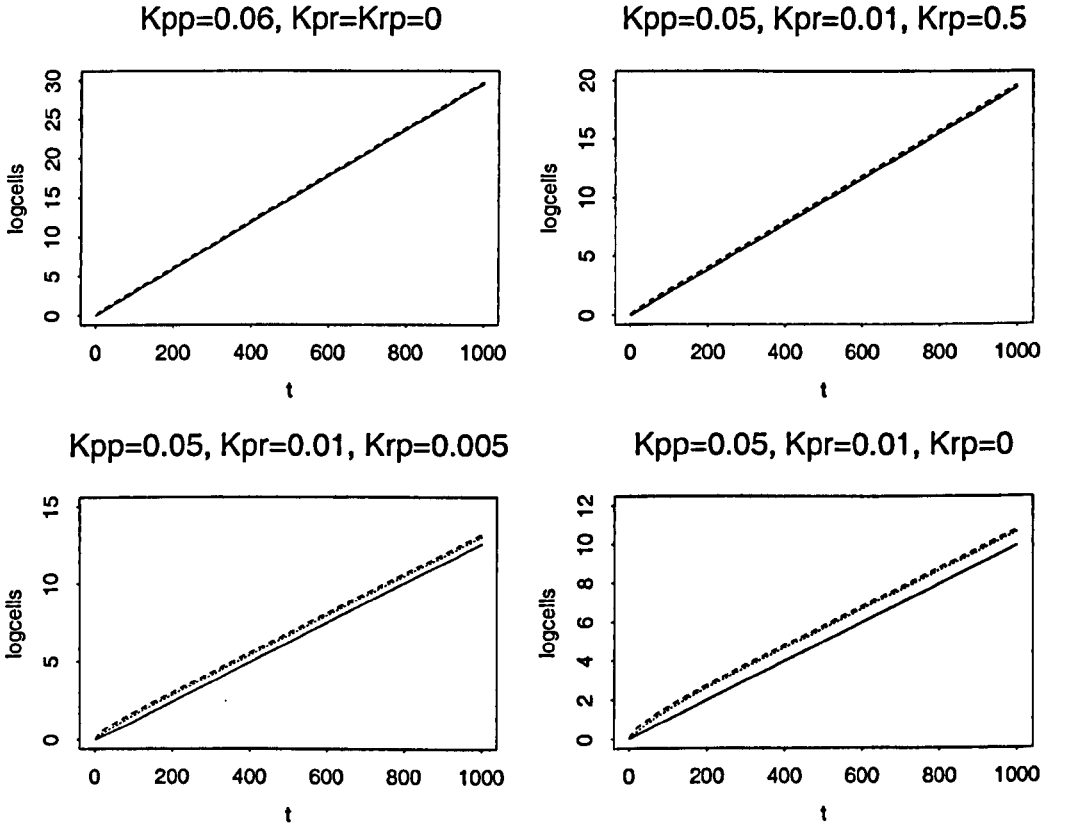


Figure 5.5: Multi-Type Birth and Death Processes. Four different processes with $k_d$ as indicated. In all cases, $k_{pp} = 0.06$, $k_{pr} = k_{rp} = 0$ and $k_{pi} = 0.05$. Cumulative plots of the expected number of cells of each type in logarithms over time, starting with one proliferative cell. Straight lines correspond to proliferative cells and dashed lines to irreversibly differentiated cells.

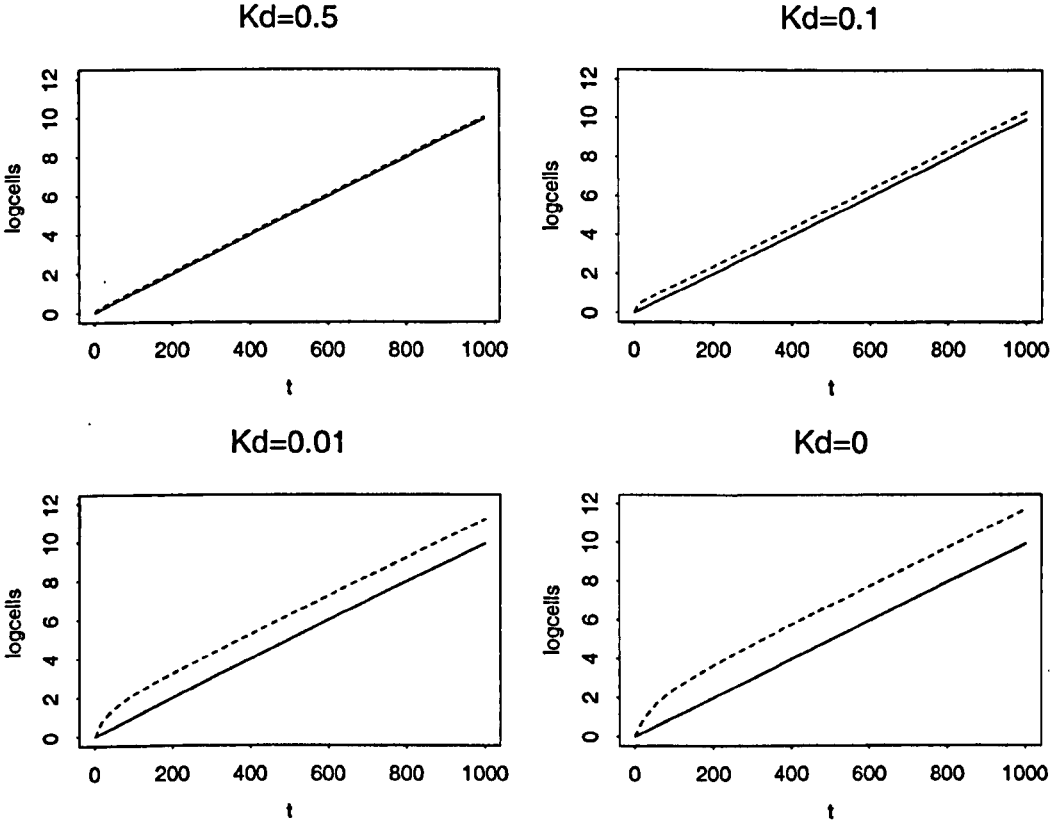


Figure 5.5 intends to illustrate that the time spent by irreversibly differentiated cells before they actually leave the organism can originate appreciable changes in the expectation of the total number of cells in the whole population. Longer times waiting for death (that is, smaller $k_d$, which would mean that cells get irreversibly differentiated in earlier phases of the cell cycle) originate bigger irreversibly differentiated supopulation without altering the pool of proliferative cells.

## 5.4 Multiple-Phase Birth and Death Processes

The cell cycle is known to be a compound of stochastic phases. All the models considered above rely on the assumption of random transitions from phase to phase imposing exponential distribution waiting times between transitions. Even though this kind of models could be realistic in many cases, some real processes demand less restrictive assumptions. Exponential generating times would not be appropriate to model the cell cycle if there is an evidence that, for instance, a new born cell requires some minimum maturation time before splitting.

The Markovian property simplifies the tractability of all the stochastic processes analysed before. This property will be lost if we consider a general distribution for the life-time, making the stochastic processes more complicated and, sometimes, intractable.

We could easily generalise the exponential generation time to a $\chi^2$ distribution with even degrees of freedom (note that the exponential distribution is a $\chi^2$ distribution with 2 degrees of freedom)by appropriately applying and generalising the model presented in Bailey (1967, pp. 131 – 135). This approach was first used by Kendall (1948).

In what follows, the life cycle is artificially divided in pseudophases (without biological counterpart) imposing transition probabilities such that transitions from phase to phase occur at random so that the Markovian property holds.

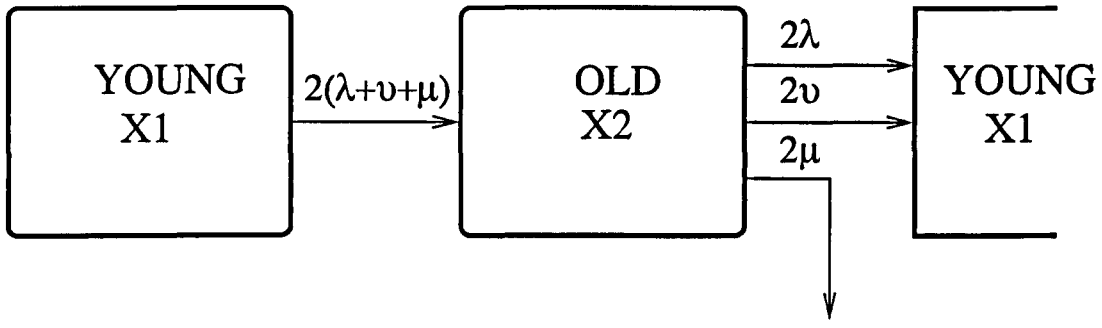
Let us imagine that the life-time of each cell consists of $k$ consecutive pseudophases. At each time $t$, $X_1(t), ..., X_k(t)$ represents the number of cells in each of these $k$ pseudophases. When transitions from phase to phase occur at random, the Markovian property holds. But, more importantly, as we will see, if the life lengths are exponentially distributed, then the overall life length will have the more general $\chi^2$ distribution with $2k$ degrees of freedom.

In Figure 5.3 the life of a cell is, artificially, divided in $k = 2$ compartments. A new born cell enters the first compartment or youth phase, $X_1$. During the period $\Delta t$, a young cell will be transferred to the second phase with probability, let us say, $2\beta\Delta t$. Hence, the distribution of stay in the first phase is the exponential distribution with parameter $2\beta$, which apart from a scale factor is essentially a $\chi^2$ distribution with two degrees of freedom. Imposing that a cell in the second or maturity phase $X_2$ has the same $2\beta\Delta t$ probability of dying, the distribution of stay will be the same as in the first phase. Consequently, as the life of the cell is divided in two exponentially distributed phases, *i.e.* two $\chi^2$ with 2 degrees of freedom, the overall life length will be distributed as a $\chi^2$ with $2 \times 2$ degrees of freedom.

The rate $\beta$ will be designed so that the model acknowledges the different behaviours of each of the three cell types described in Section 5.1).

We assume that a young cell cannot die and that will eventually become mature. It can then choose to go on as a proliferative cell, to go resting or to leave the cycle irreversibly. The decision that cells take will determine the fate of the population. If the dead cell is replaced by two new born cells, the cell is proliferative. If it is replaced by just one new born cell, the cell decides not

Figure 5.6: Cell Cycle for the Multiple-Phase Birth and Death Process. $2(\lambda + v + \mu)$ is the ratio of young cells that become mature and $2\lambda$, $2v$ and $2\mu$ are the ratio of mature cells that split into two daughter cells, the ratio of cells that go resting and the ratio of cells that die respectively.




to follow the cycle but to rest. Finally, if there is not any new cell, then the cell chooses to die.

Therefore, in this process during $\Delta t$ there are four possible types of changes in the subpopulations with the following probabilities:

- Prob. $X_1$ increases by two and $X_2$ decreases by one (cell division): $2\lambda X_2(t)\Delta t + o(\Delta t)$.
- Prob. $X_1$ increases by one and $X_2$ decreases by one (a proliferative cell goes resting): $2vX_2(t)\Delta t + o(\Delta t)$.
- Prob. $X_1$ does not change and $X_2$ decreases by one (cell death): $2\mu X_2(t)\Delta t + o(\Delta t)$.
- Prob. $X_1$ decreases by one and $X_2$ increases by one (a cell becomes mature): $2(\lambda + v + \mu)X_1\Delta t + o(\Delta t)$.

Note that the last probability has been assigned so that the life length distributions of both subpopulations are the same, *i.e.* exponential with parameter $2\beta$, where $\beta = \lambda + v + \mu$. Consequently, the time spent from birth to death

has essentially a $\chi^2$ distribution with 4 degrees of freedom, and expected value $1/\beta$.

By an straightforward extension of equation (5.27), the corresponding differential equation for the joint cumulant-generating function can be written as

$$\frac{\partial K}{\partial t} = 2(\lambda+v+\mu)[e^{-\theta_1+\theta_2}-1]\frac{\partial K}{\partial \theta_1}$$

$$+(2\lambda[e^{2\theta_1-\theta_2}-1]+2v[e^{\theta_1-\theta_2}-1]+2\mu[e^{-\theta_2}-1])\frac{\partial K}{\partial \theta_2}, \qquad (5.59)$$

where the same notation as in previous sections is used.

Equating coefficients on both sides of equation (5.59), we obtain the following set of partial differential equations for the first order cumulants:

$$\frac{\partial \kappa_{10}}{\partial t} = -2(\lambda+v+\mu)\kappa_{10}+2(2\lambda+v)\kappa_{01}, \qquad (5.60)$$

$$\frac{\partial \kappa_{01}}{\partial t} = 2(\lambda+v+\mu)\kappa_{10}-2(\lambda+v+\mu)\kappa_{01}. \qquad (5.61)$$

$$(5.62)$$

Solving this system gives the first order moments of $X_1(t)$ and $X_2(t)$:

$$\kappa_{10}(t) = \frac{1}{2}e^{-2(\beta-\sqrt{\varphi\beta})t}+\frac{1}{2}e^{-2(\beta+\sqrt{\varphi\beta})t} \qquad (5.63)$$

$$\kappa_{01}(t) = \sqrt{\frac{\beta}{\varphi}}(\frac{1}{2}e^{-2(\beta-\sqrt{\varphi\beta})t}-\frac{1}{2}e^{-2(\beta+\sqrt{\varphi\beta})t}) \qquad (5.64)$$

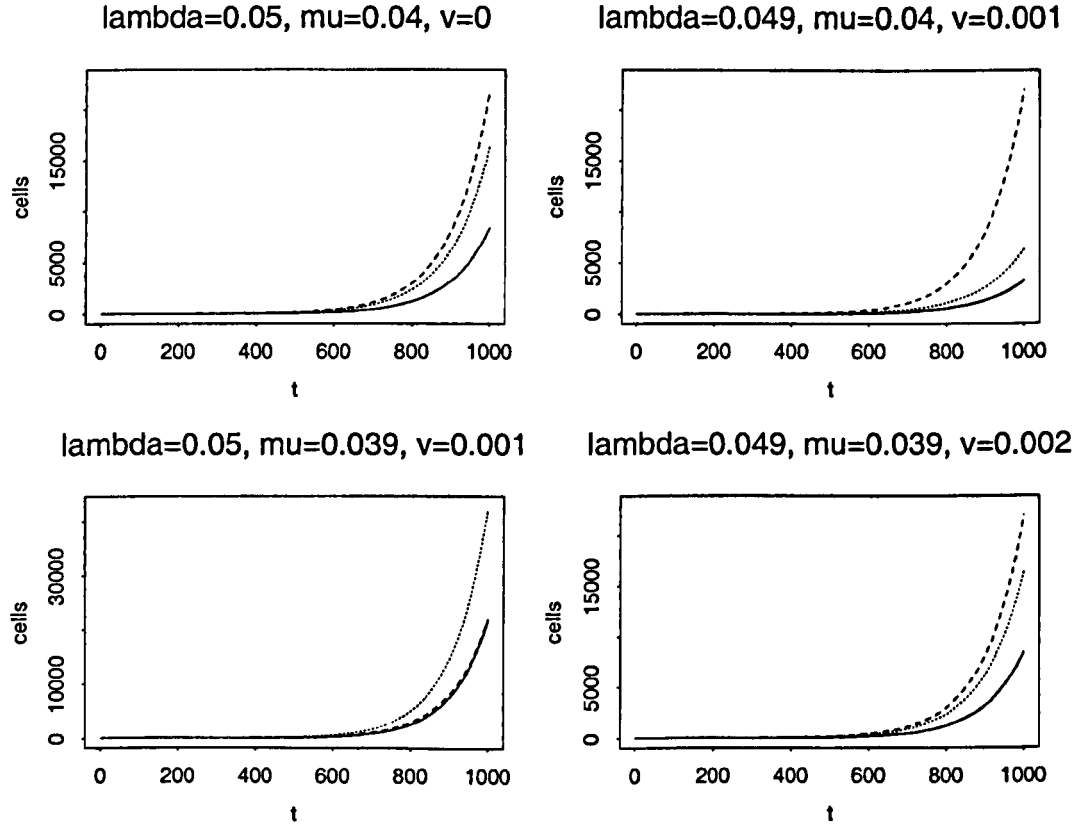
where $\beta=\lambda+v+\mu$ and $\varphi=2\lambda+v$. Therefore, the expected total number of cells at time $t$ will be given by:

$$m(t) = \kappa_{10}(t)+\kappa_{01}(t) \qquad (5.65)$$

$$= e^{-2(\beta-\sqrt{\varphi\beta})t}\frac{\sqrt{\varphi}+\sqrt{\beta}}{2\sqrt{\varphi}}+e^{-2(\beta+\sqrt{\varphi\beta})t}\frac{\sqrt{\varphi}-\sqrt{\beta}}{2\sqrt{\varphi}} \qquad (5.66)$$

In figure 5.7 the multiple-phase birth and death process is compared to the simple birth and death process with parameters $\lambda=0.05$ and $\mu=0.04$ (therefore, with expected life length of 11.11). The expected number of cells of the

Figure 5.7: Multiple-Phase and Standard Birth and Death Processes. Dashed lines represent the expected number of cells in a simple birth and death process with $\lambda = 0.05$ and $\mu = 0.04$. The expected number of cells in a multiple-phase process with the indicated instant probabilities are represented by dotted lines, where straight lines account for the corresponding subpopulation of young cells. In all cases the populations start with only one cell.



simple birth and death process in all four plots by dashed lines. In all the cases the expected life length of each cell is fixed at 11.11 with different instant probabilities. Straight lines represent the expected number of young cells, whereas dotted lines account for the total multiple-phase population.

A comparison of the plots in figure 5.7 shows that the rate of cells that rest plays a vital role in determining the evolution of the expected size of the population.

## 5.5 Branching Processes

### 5.5.1 Introduction

In this section another kind of stochastic processes will be introduced: Branching Processes. Bellman and Harris (1952) first introduced and studied the theory of Branching Processes and an extensive literature has followed (see general references as, for instance, Harris, 1963, Athreya and Ney, 1972, and Jagers, 1975, and references therein; and Christensen and Shonkweiller, 1978, Cowan, 1985, Cowan and Morris, 1998, and Jagers and Klebaner, 2000, for more recent developments).

### 5.5.2 The General Model

We are interested on a population with one ancestor composed by individual cells, $x$, characterised by the pairs $(\lambda_x, \xi_x)$, where $\lambda_x$ is the life length of $x$ and $\xi_x$ the reproduction point process of $x$, which are supposed to be *i.i.d* [2] with probability distribution $Q$, known as the *individual law*, measured on the space $R_+ \times N(R_+)$.

[2]Therefore, we will omit subscript $x$.

These kinds of processes are called *Branching Processes.* For a full summary of the theory see, for instance, the books by Athreya and Ney (1972) and Jagers (1975). We shall follow the notation and conventions of the latter one, and details and the proofs of the theorems should be consulted there.

The margin of Q on $R_+$ is the *life-length distribution* and the margin on $N(R_+)$ the *reproduction law.* If $\sigma_x$ is the *birth time* of $x$ and $a \geq 0$, the variable

$$z_t^a(x) = \begin{cases} 1 & \text{if } t - a < \sigma_x \leq t < \sigma_x + \lambda_x \\ 0 & \text{otherwise} \end{cases}$$

will be one if $x$ is alive and is younger than $a$ at time $t$. Therefore,

$$z_t^a = \sum_x z_t^a(x) \leq \infty \tag{5.67}$$

counts all the individuals younger than $a$ that are alive at time $t$. The stochastic process $\{z_t^a, t, a \geq 0\}$ is the *general branching process.* When, there is no interest on the age distribution of the cells, *e.g.* $a = \infty$, we will write $\{z_t, t \geq 0\}$. Similarly, if $y_t(x)$ takes value one when $x$ has been born (*i.e.*, $\sigma_x \leq t$),

$$y_t = \sum_x y_t(x) \tag{5.68}$$

will be the total number of individuals that have been born up to time $t$.

Some more notation will be needed in what follows. The life-length distribution function and the expected value of the process will be denoted as $L(u) = P[\lambda \leq u]$, and $m_t^a = E[z_t^a]$ respectively. The *reproduction generating function* will be given by:

$$f(s) = E[s^{\xi(\infty)}] \tag{5.69}$$

and we define the *reproduction function* as, $\mu(t) = E[\xi(t)]$, and the *reproduction mean* as $m = f'(1) = \mu(\infty)$.

We can classify the $z(t), t \geq 0$ according to the corresponding reproduction mean, $m$, as: *supercritical* ( $m > 1$ ), *critical*: ( $m = 1$ ) or *subcritical*: ( $m < 1$ ).

On the other hand, a process is called *Malthusian* when there is a number $\alpha$ (the *e.g. Malthusian* parameter) such that:

$$\widehat{\mu}(\alpha) = \int_0^\infty e^{-\alpha t}\mu(dt) = 1. \tag{5.70}$$

The Malthusian parameter will be positive, zero or negative according as $m >$ , $=$, or $< 1$ respectively. The Malthusian parameter will always exist for supercritical processes. This will not always be true in the subcritical cases.

The process will be *lattice* if there is a number $d > 0$ such that

$$\sum_{k=0}^{\infty} \mu(\{kd\}) = \mu(\infty). \tag{5.71}$$

We can now state some important theorems. Theorems 5.1 and 5.2 are taken from the book by Jagers (1975, Theorems 6.5.1 and 6.3.3 [3]).

**Theorem 5.1.**

*If the reproduction function is finite, then so is* $E[y_t]$ *and therefore also* $m_t = E[z_t]$ *for all* $t$. *Further,* $m_t^a = E[z_t^a]$ *satisfies*

$$m_t^a = 1_{[0,a)}(t)\{1 - L(t)\} + \int_0^t m_{t-u}^a \mu(du).$$

*If* $m = \mu(\infty) < 1$ *(the subcritical case), then as* $t \to \infty$ *then* $m_t \to 0$.

*If* $m = 1$ *(the critical case) and* $\mu$ *is non-lattice, then for* $0 \leq a < \infty$

$$m_t^a \to \frac{\int_0^a \{1 - L(u)\}du}{\int_0^\infty u\mu(du)}$$

*If further,*

$$\int_0^\infty tL(dt) < \infty,$$

[3] There are obvious lattice analogs.

*then also*

$$m_t \to \frac{\int_0^\infty uL(du)}{\int_0^\infty u\mu(du)}.$$

*When $m > 1$ (the supercritical case), $\mu$ is not lattice, and $\alpha > 0$ is the Malthusian parameter defined by $\widehat{\mu}(\alpha) = 1$, then for $0 \leq a \leq \infty$,*

$$m_t^a \sim e^{\alpha t}\frac{\int_0^a e^{-\alpha u}\{1 - L(u)\}du}{\int_0^\infty ue^{-\alpha u}\mu(du)}.$$

*In the lattice cases corresponding assertions hold.*

**Theorem 5.2.**

*In a non-lattice, subcritical process admitting the Malthusian parameter $\alpha$,*

$$m_t^a \sim e^{\alpha t}\frac{\int_0^a e^{-\alpha t}\{1 - L(t)\}dt}{\int_0^\infty te^{-\alpha t}\mu(dt)} \tag{5.72}$$

*for $0 \leq a < \infty$, as $t \to \infty$. For $a = \infty$ the relation still holds, provided*

$$\int_0^\infty te^{-\alpha t}L(dt) < \infty. \tag{5.73}$$

Therefore, we have that under certain conditions, the total expected number of cells of all ages in a population, $m_t^\infty = m_t$, will be given by:

$$m_t \sim e^{\alpha t}\frac{\int_0^\infty e^{-\alpha t}\{1 - L(t)\}dt}{\int_0^\infty te^{-\alpha t}\mu(dt)} = e^{\alpha t}k(t). \tag{5.74}$$

A particular case of the general model will be now described in more detail, the *Bellman-Harris Process*, suitable to study the evolution of cell populations.

### 5.5.3 Bellman-Harris Processes

Splitting processes or *Sevast' yanov* processes were introduced by Sevast' yanov (1964) and the *Bellman-Harris process* processes are known since the forties (Bellman and Harris, 1952).

**Definition 5.1.**

*When $P(\xi\{\lambda\} = \xi\{\infty\}) = 1$, the process is called splitting or Sevast' yanov process, and if in addition, $\lambda$ and $\xi$ are independent, is known as the 'Bellman-Harris' or the 'age dependent' process.*

In other words, in a *splitting* process, new cells are only born at the death of their mother. If further, the reproduction process is independent of the life-span we will talk about *Bellman-Harris* processes.

We are concerned with a stochastic model for the cell proliferation. As discussed above, at the end of their lifes cells can either die or split no matter the length life (note that cells that have been resting when back in the cycle they become proliferative again and, therefore, when they reach mitosis can either die or split).

Therefore, given the characteristics of cells' behaviour, the stochastic process that we will construct for the population will be a *Bellman-Harris* process. This fact will simplify the reproduction function of the process to:

$$\mu(u) = \int_0^t f'_u(1)L(du) = \int_0^t m(u)L(du) = mL(u). \tag{5.75}$$

From equation (5.70), the Malthusian parameter $\alpha$ satisfies,

$$1 = \widehat{\mu}(\alpha) = \int_0^\infty e^{-\alpha y} mL(dy) = m\int_0^\infty e^{-\alpha y} l(y)dy, \tag{5.76}$$

and applying logarithms it follows that,

$$-\ln m = \ln \int e^{-\alpha y} l(y)dy. \tag{5.77}$$

Noting that $\ln \int e^{-\alpha y} l(y)dy = K(-\theta)$, where $K$ is the cumulant generating function, using the Taylor expansion we get:

$$\ln \int e^{-\alpha y} l(y)dy = \sum_{r=1} \kappa_r \frac{(-\theta)^r}{r!} = -\alpha\mu + \frac{\alpha^2}{2\sigma^2} + O(\alpha^3), \tag{5.78}$$

where $\mu$ and $\sigma^2$ denote the mean and variance of $l(y)$ respectively. Setting $m = 1 + \epsilon$ and solving for $\alpha$ in terms of $\epsilon$, it follows that

$$\alpha = \frac{\epsilon}{\mu} + \frac{\epsilon^2}{2\mu}(\frac{\sigma^2}{\mu^2} - 1) + O(\epsilon^3). \tag{5.79}$$

Empirically, restricting our attention to first and second order moments, and after testing multiple expressions, the best approximation for the Malthusian parameter for the distributions we have considered is the following one:

$$\alpha \simeq \frac{\mu}{\sigma^2}((m)^{\sigma^2/\mu^2} - 1) = \widehat{\alpha}. \tag{5.80}$$

Directly expanding equation (5.76), setting $m = 1 + \epsilon$, it follows immediately equation (5.80), where the error will be of $O(\epsilon^3)$ if the third moment is finite.

On the other hand, recalling equation (5.74), for the *Bellman-Harris* process we have that:

$$k(y) = \frac{\int e^{-\alpha y}(1 - L(y))dy}{m \int y e^{-\alpha y} l(y) dy}. \tag{5.81}$$

Now, differentiating with respect to $m$ in equation (5.76),

$$\frac{\partial \alpha}{\partial m} = \frac{1}{m^2 \int y e^{-\alpha y} l(y) dy}, \tag{5.82}$$

and therefore, it follows that

$$\int_0^\infty y e^{-\alpha y} l(y) d(y) = \frac{1}{m^2 \frac{\partial \alpha}{\partial m}} \simeq \frac{\mu}{m m^{\frac{\sigma^2}{\mu^2}}}, \tag{5.83}$$

where the last approximation is due to equation (5.80).

On the other hand,

$$\int e^{-\alpha y}(1 - L(y))dy = -\frac{1}{\alpha}\int (1 - L(y))d(e^{-\alpha y}) \tag{5.84}$$

$$= [-\frac{1}{\alpha}(1 - L(y))e^{-\alpha y}]_0^\infty + \frac{1}{\alpha}\int e^{-\alpha y} d(1 - L(y)) \tag{5.85}$$

$$= \frac{1}{\alpha} - \frac{1}{\alpha}\int e^{-\alpha y} l(y) dy = \frac{1}{\alpha}(1 - \frac{1}{m}) = \frac{m-1}{\alpha m}. \tag{5.86}$$

Consequently, $k$ can be approximated as:

$$k \simeq \frac{m-1}{\widehat{\alpha} m} \frac{m^{\frac{\sigma^2}{\mu^2}}}{\mu} = \widehat{k}, \tag{5.87}$$

and therefore, combining the approximations to the malthusian parameter and to the constant $k$, it follows that:

$$m_t \simeq \widehat{m}_t = \widehat{k} e^{\widehat{\alpha} t} \tag{5.88}$$

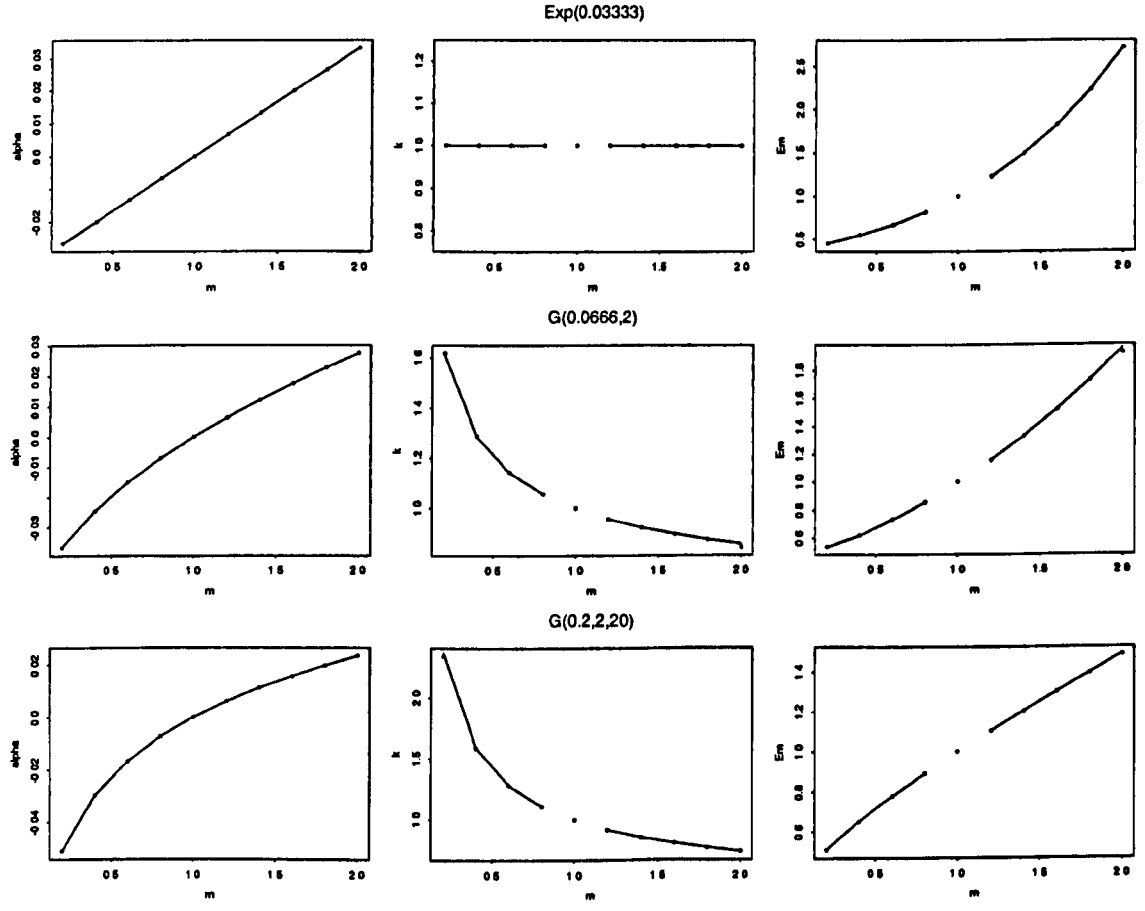
This expression constitutes a simple approximation to the expected number of cells alive at time $t$. The main advantage of this result is that it does not depend on the particular life distribution of cells. The first and second order moments fully characterise this expectation.

If $p \in [0,1]$ denotes the probability of cell division, the reproduction mean of the process will be given by $m = 2p \in [0,2]$. Dots in figure 5.8 represent the real values of $\alpha$, $k$ and $m_t$ at time $t = 30$ for ten different values of the reproduction mean $m$. Three different life length distributions are considered:

- Exponential distribution with parameter 0.03333.
- Gamma distribution with parameters 0.06666 and 2 with a *minimum life length* of 2 (that is to say that the probability of dividing or splitting is zero in the first 2 time units of life of cells).
- Gamma distribution with parameters 0.2 and 2 with a minimum *life length* of 20.

In all cases the approximations given by equations (5.80), (5.87) and (5.88) are very close to the real values (note that in the exponential case the approximations equal the real values). This simplifies enormously the calculation of the expected number of cells in a population for large populations. Information about the mean and the variance of the distribution of cell life lengths comprises enough information to characterise the limit result for the expected number of cells.

Figure 5.8: Approximations. Real values of the malthusian parameter, $\alpha$, the constant $k$ and the expected number of cells in the population, $m_t$, at time $t = 30$ are represented by dots. Three different life length distributions are considered, one exponential and two gamma distributions with a minimum time length for either division or death. The obtained approximations are represented by straight lines.



## 5.6 Conclusions

In this chapter we have studied the kinetics that govern the proliferation of cell populations. After introducing the cell cycle, we have presented several deterministic and stochastic models.

Two different stochastic processes have been considered in this chapter in order to build models for cell growth: 'continuous time Markov processes' and 'Branching processes'. The underlying assumptions and the pros and cons of each of these processes have been detailed and discussed. Three novel models have been derived: the 'Multi-Type Birth and Death Process', the 'Multiple-Phase Birth and Death Process' and the 'General Life Length Process'. Without a particular data set or a research interest in mind, these three models should be considered as complementary as both the underlying assumptions and the degree of detail of results are different in each of them.

In what follows, cell populations will be exposed to drugs that alter their kinetic behaviour. Therefore, drug-effect will be introduced into the cell proliferation process.

# Chapter 6

# Pharmacodynamics

In chapters 2 and 3 various models to estimate the concentration of drugs in different body compartments containing certain cell, tissues and/or organs, have been introduced. The next step is to try to quantify, and therefore understand, the effects that a certain amount of drug induce within the body during treatment.

In the first sections of this chapter we are going to present a brief summary of some of the most widely used pharmacodynamic models linking plasma drug concentration with response. Our main interest lies on analysing the fate of cell populations which are exposed to drugs. The aim is to both qualitatively and quantitatively describe the effect that certain blood drug concentration levels have in the evolution of cell populations.

In section 6.2.3, we construct a new general model that links pharmacokinetic, pharmacodynamic and cell kinetics processes. An additional crucial phenomenon will be introduced in the model in section 6.2.4: 'drug resistance'. It is well known that treatment-resistant cells can arise at any time during treatment. These cells can in time come to comprise subpopulations of cells that will not be affected by the original treatment.

We are going to perform toy examples in order to see how the models we propose in this chapter behave in a deterministic world, free of measurement error. The evolution of leukaemic cell populations will be analysed under four different treatment scenarios.

The population approach, already introduced in chapter 3 to deal with pharmacokinetic data, will be then adopted to, hopefully, handle the variability that in most cases clinical data present.

## 6.1 Introduction

Pharmacodynamics is the study of the biochemical and physiological effects of drugs and their mechanisms of action. It studies the interactions between plasma drug concentration and the caused effects in the targed cells, tissues and/or organs.

Physiological effects of drugs can be divided into two groups: 'continuous' and 'quantal' or 'dichotomous'. A continuous response can be evaluated as a continuous function of dose, concentration and time (*e.g.* blood pressure and heart rate). A quantal effect, on the other hand, is a discrete variable, there is either a response or not, nothing in between (*e.g.* death and remission). In what follows, we are going to concentrate on continuous effects.

In addition, the type of relationship between the plasma drug concentration and a given response can generally be determined by two factors: whether concentration is directly or indirectly related to response, and whether the drug interacts with the receptor reversibly or irreversibly.

Lastly, the relationship between effect and concentration can be either 'agonistic' or 'antagonistic'. The effect is said to be agonistic when its magnitude

increases with higher drug concentrations and antagonistic when there is an inverse relationship between the magnitude of the effect and the levels of concentration. In what follows, we will only consider agonistic effects.

There are innumerable specific models for the interactions between particular drugs and the different effects they cause. Ross (1996) gives account of most of the pharmacodynamic models that have been used in clinical applications and research. Particular models can be found in Levy (1966), Jusko (1971), Matthews (1993), Jusco and Ko (1994), Levy (1994) and Jenkinson *et al* (1995). An exploratory analysis of the data set will in most cases provide hints about the most adequate model in each case.

## 6.2 Main Pharmacodynamic Models

### 6.2.1 Introduction

It has already been mentioned that the type of relationship between the plasma drug concentration and a given response can be either direct or indirect. In the first case the intensity of a given pharmacological response is due to a direct effect of the drug on the receptor. When the relationship is indirect, the intensity of the response may be the net result of several processes, only one of which is influenced by the drug. If this is the case, the process that is influenced by the drug must be identified and an attempt made to relate plasma drug concentrations to changes in this process.

## 6.2.2 Direct Pharmacodynamic Models

### Introduction

In a direct pharmacodynamic model, by definition, the drug concentration at the receptor site will be a function of the drug concentration in the plasma. The form of the relationship will depend on the location of the site of action. Following the compartmental approach taken in chapter 2 the site of action may be located either in the central compartment or in one of the peripheral compartments.

The site of action may be determined by examining the relationship between the intensity of response and the concentration of drug in the plasma or the estimated amount of drug in the corresponding peripheral compartment.

### One-Compartment Models

When the drug concentration over time in the plasma can be described by a one-compartment model, the drug concentration at the receptor site, $C_r$, will be proportional to the drug concentration in the plasma, $C$, as the receptor is believed to be located in the same compartment as the circulatory system.

The simplest relationship between concentration and an agonistic pharmacological effect, $E$, can be described by a linear function of the form,

$$E(t) = E_0 + qC(t), \tag{6.1}$$

where the intercept, $E_0$, and the slope, $q$, are constants to be determined.

However, a linear and direct proportionality between concentration and effect rarely applies to actual cases. Nevertheless, in a number of instances, a linear proportionality between the effect and the log-transformation of the

concentration has been applied successfully:

$$E(t) = E_0' + q' \ln C(t), \tag{6.2}$$

where the intercept $E_0'$ and the slope $q'$, are constants to be determined. In this model, the pharmacologic response can not be estimated when the concentration is zero because of the logarithmic function.

In addition, in both (6.1) and (6.2) the effect tends to infinity as the concentration increases. This violates a widely accepted pharmacologic principle which under certain biological conditions states that there is a drug-induced 'maximal effect'.

A widely used relationship between concentration and an agonistic effect has the following form:

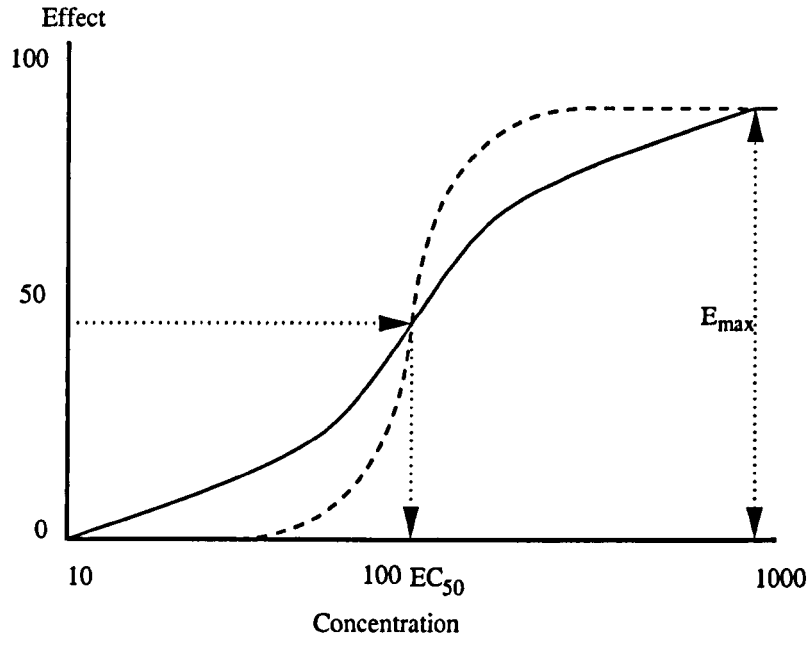
$$E(t) = \frac{E_{\mathrm{MAX}} C(t)^m}{EC_{50}^m + C(t)^m}, \tag{6.3}$$

where $E_{\mathrm{MAX}}$ is the drug induced maximum effect and $EC_{50}$ represents the 'potency' of the drug, and $m > 0$ is known as the 'sigmoidicity factor'. A baseline effect could be added if necessary.

Potency is an expression of the activity of a compound in terms of the concentration needed to produce a defined effect. The variable $EC_{50}$ is the concentration of the drug that produces 50% of the maximal possible effect. The sigmoidicity parameter $m$ does not necessarily have a direct biological interpretation but it provides a degree of flexibility in the sensitivity of the response-concentration relationship. This equation is known as the *Sigmoid* $E_{\mathrm{MAX}}$ Model or the *Hill equation.* This equation was first used by Hill (1910) and introduced in pharmacodynamics by Wagner (1968).

The Hill equation will quantitatively and fully characterise the typical sigmoid curve resulting from a effect versus concentration plot on a logarithmic

Figure 6.1: Hill Equation.The dashed line represents the Hill equations with a higher sigmoidicity factor, with same potency and maximal effect, than the solid line.




concentration scale as shown in figure 6.1. The solid line has a small sigmoidicity factor, whereas the dashed line is the result of a higher sigmoidicity factor. Both lines have the same potency and maximal effect. A high exponent results in an all-or-nothing effect.

When treatment consists in more than one drug, the most common generalisation of the Hill equation (see, for example, Gero, 1971, and Ebling *et al*, 1991) is given by:

$$E(t) = \sum_{l} \frac{E_{\mathrm{MAX},l} C_l(t)^{m_l}}{EC_{50,l}^{m} + C_l(t)^{m_l}}, \tag{6.4}$$

where $l = 1, ..., L$, $L$ being the number of different drugs administered. This model assumes that there are no significant interactions between drugs and that the effect they all produce is agonistic. For interaction pharmacodynamic models, see for instance, Kenakin (1993) and Matthews (1993).

### Multi-Compartment Models

The time course of drug action in multicompartment systems depends on the location of the site of action. When the site of action is associated with the central compartment, a plot of effect versus the log-concentration should yield the same sigmoid-type curve as that shown in figure 6.1.

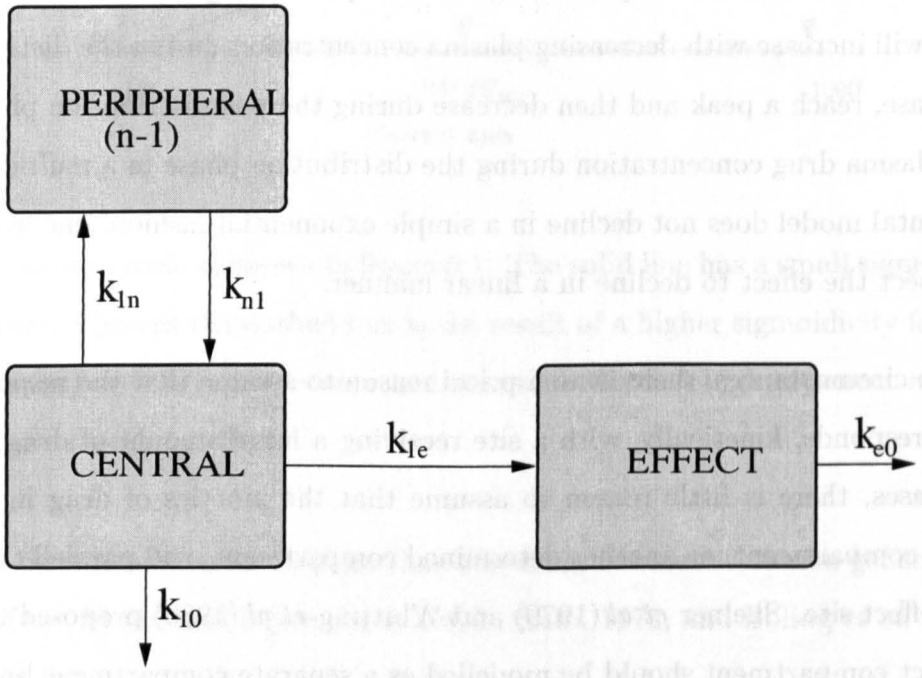
A similar relationship should also be observed when the effect is associated with the concentration in a peripheral compartment when the effect is plotted against the estimated log-concentration of the corresponding peripheral compartment.

However, when the effect is plotted against the plasma log-concentration, response will increase with decreasing plasma concentration during the distribution phase, reach a peak and then decrease during the postdistribution phase. Since plasma drug concentration during the distribution phase in a multicompartmental model does not decline in a simple exponential fashion, one would not expect the effect to decline in a linear manner.

In some circumstances, there is no a priori reason to assume that the response site corresponds, kinetically, with a site receiving a large amount of drug. In these cases, there is little reason to assume that the kinetics of drug in the plasma compartment, or another determined compartment, will paralell those at the effect site. Sheiner *et al* (1979) and Whitting *et al* (1980) proposed that the effect compartment should be modelled as a separate compartment linked to the plasma compartment by, for instance, a first-order process receiving a negligible amount of drug. Consequently, there is no need to add an additional exponential term into the n-compartmental pharmacokinetic model to account for the effect compartment.

The model presented by Sheiner *et al* (1979) is illustrated in Figure 6.2. The rate constants from the central to the $n-1$ peripheral compartments have

Figure 6.2: Multi-compartment Model. There is a central compartment from where drug is delivered into all other compartments, including the effect compartment, at rates $k_{1n}$ and $k_{1e}$ respectively. Subtances return to the central compartment from the peripheral compartments at rates $k_{n1}$ and is eliminated at rate $k_{10}$. It is assumed that elimination occurs directly from the effect compartment at rate $k_{e0}$.



already be introduced in Chapter 2. The novelty of this model is that a first order rate constant $k_{1e}$ connects the central to the effect compartment. If we assume $k_{1e}$ to be very small relative to the magnitude of any other rate constant in the model, the drug transfer from the central compartment to the effect compartment is negligible, and therefore does not influence the plasma concentration versus time plot. As the amount of drug in the effect compartment will be negligible, it can be eliminated directly from the effect compartment, with rate constant,say, $k_{e0}$.

The concentration of drugs in the different compartments can easily be obtained following the techniques presented in chapter 2 (see, for instance, Sheiner *et al*, 1979). For a two-compartment model ($n = 2$ in Figure 6.2), the concentration levels in the central and effect compartment following a single intravenous administration are given by:

$$C_c(t) = \frac{D}{V_1}\left(\frac{(k_{12}-\alpha)e^{-\alpha t}}{\beta-\alpha} + \frac{(k_{12}-\beta)e^{-\beta t}}{\alpha-\beta}\right), \tag{6.5}$$

$$\begin{aligned} C_E(t) = &\frac{k_{e0}D(k_{12}-\alpha)}{V_1(\beta-\alpha)(k_{e0}-\alpha)}e^{-\alpha t} + \frac{k_{e0}D(k_{21}-\beta)}{V_1(\beta-\alpha)(k_{e0}-\beta)}e^{-\beta t} + \\ &+ \frac{k_{e0}D(k_{12}-k_{e0})}{V_1(\alpha-k_{e0})(\beta-k_{e0})}e^{-k_{e0}t}, \end{aligned} \tag{6.6}$$

where $V_1$ is the volume of distribution in the central compartment, $D$ the administered intravenous dose, and $\alpha$ and $\beta$ are given by:

$$\begin{aligned} \beta &= \frac{1}{2}\left((k_{12}+k_{21}+k_{10}) - \sqrt{(k_{12}+k_{21}+k_{10})^2 - 4k_{21}k_{10}}\right), \\ \alpha &= \frac{k_{21}k_{10}}{\beta}. \end{aligned}$$

Therefore, we have obtained an expression for the amount of drug reaching the effect compartment during treatment. This fraction of the total dose will determine the effect of interest. An estimation of the substance that actually enters into the effect compartment is of great importance when choosing optimal doses.

### 6.2.3 Indirect Pharmacodynamic Models

#### Introduction

The intensity of a pharmacologic response may not be caused by a direct effect of the drug but it may be the net result of several processes, only one of which is influenced by the drug. Under such circumstances a direct relationship between the plasma concentration of the drug and the effect can not be obtained.

This class of models is typically based on some biological principles with prior knowledge about the mechanisms of drug action in that the variables and parameters have a physiological meaning. In contrast, the previous direct models rely on empirical concepts.

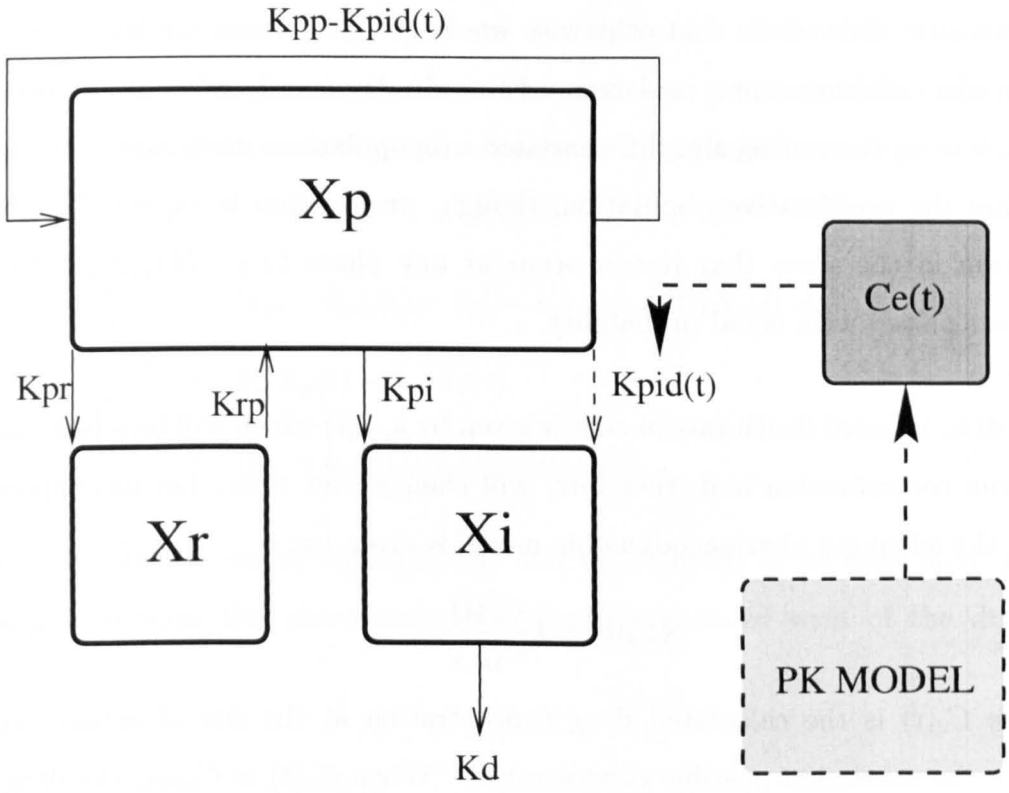
The particular process that is altered by the drug must first be identified and then the relationship between plasma drug concentration and this process must be modelled. Applications of indirect pharmacodynamic models can be found, for instance, in Jusco (1971), Jusco (1990) and Dayneka *et al* (1993).

Our interest will lie on the cell kinetic processes analysed in the previous chapter. The effect that drugs will originate in the cell cycle need to be modelled in order to establish the effect that treatment has in the evolution of cell populations.

#### Cell Death

Although most drugs produce a response that is reversible, certain antibiotics and anticancer agents cause cell death, an irreversible effect, by attacking the cell cycle. It is appropriate to classify two different types of drugs, each of which affects the cell cycle in a different manner: cell cycle phase-specific and non-phase-specific drugs. Each type requires a different model as the affected population of cells differs from one type to the other.

Figure 6.3: Pharmacokinetic-Pharmacodynamic-Cell-Kinetics Model. There are three types of cells, the proliferative cells, $X_p$, the resting cells, $X_r$, and the irreversibly differentiated cells, $X_i$, with migration rates $k_{pr}$,$k_{rp}$, $k_{pi}$, $k_d$ and splitting rate $k_{pp}$. When treatment starts, the pharmacokinetic model will determine the concentration of drugs in the effect compartment, $C_e(t)$, that will have an effect on the cell population by altering the cell cycle. In particular, we assume that a fraction $k_{pid}(t)$ of the proliferative cells will not divide during mitosis as a consequence of treatment.




In what follows, we are going to introduce phase-specific drugs into the deterministic cell kinetic model introduced in chapter 5. Keep in mind that the results in this model equal the expected values for the corresponding stochastic model that assumes exponential life times in each phase of the cell cycle. We are going to link the pharmacokinetic and pharmacodynami models with the cell population model. A global view of all the processes involved in determining the effect of treatment will be therefore constructed.

As discussed in chapter 5, depending on the role they undertake within the

cycle, there are three different types of cells: proliferative, resting and differentiated cells. In figure 6.3, these three different types are represented with the corresponding deterministic flows between them. The notation is the same as the one used in chapter 5. The model is the same as represented in figure 5.2 with the only difference that drugs are interfering with the cell cycle.

In this particular model we propose, drugs have effect on the death rate of proliferative cells. Cells that otherwise would go into mitosis are being killed as an effect of drug action. It is assumed that drugs can only affect proliferative cells, leaving the resting and differentiated subpopulations unchanged directly. Within the proliferative population, though, drug action is supposed to be uniform in the sense that it can occur at any phase (*e.g.* $M1$, $S$, $M2$ or *mitosis* phase) with equal probability.

The drug-induced death rate of cells is given by $k_{pid}(t)$ which will be a function of drug concentration and, therefore, will change over time. Let us suppose that the adequate pharmacodynamic model is given by:

$$k_{pid}(t) = k\frac{C_e^2(t)}{C_{\text{MAX}}^2}, \tag{6.7}$$

where $C_e(t)$ is the calculated drug concentration at the site of action and $C_{\text{MAX}}$ the maximum possible concentration. When $C_e(t) = C_{\text{MAX}}$, the drug-induced death rate, $k_{pid}$, will be equal to $k$, which accounts for the maximum response. As discussed earlier, the site of action could either be in one of the pharmacokinetic compartments or in a separate effect compartment.

This particular pharmacodynamic model that we are proposing here will obviously be adequate only for certain drugs and when looking at certain effects. An exploratory analysis should be done to determine the most adequate pharmacodynamic model to link particular drug concentration levels and a chosen response. However, the global approach presented here does not rely on any particular pharmacodynamic relationship.

When more than one drug is administered, we could easily generalise equation (6.7), so that,

$$k_{pid}(t) = \sum_{l=1}^{L} k_l \frac{C_{e_l}^2(t)}{C_{\text{MAX},l}^2}, \tag{6.8}$$

where $L$ is the number of different drugs that constitute treatment.

It follows that when treatment starts, the dynamics of the cell population changes as concentration levels vary. The model we propose, based on equations (3.2) and (6.5) is consequently determined by the following set of equations:

$$\frac{dX_p(t)}{dt} = (k_{pp} - k_{pid}(t) - k_{pr} - k_{pi} - k_{pid}(t))X_p(t) + k_{rp}X_r(t), \tag{6.9}$$

$$\frac{dX_r(t)}{dt} = k_{pr}X_p(t) - k_{rp}X_r(t), \tag{6.10}$$

$$\frac{dX_i(t)}{dt} = (k_{pi} + k_{pid}(t))X_p(t) - k_d X_i(t), \tag{6.11}$$

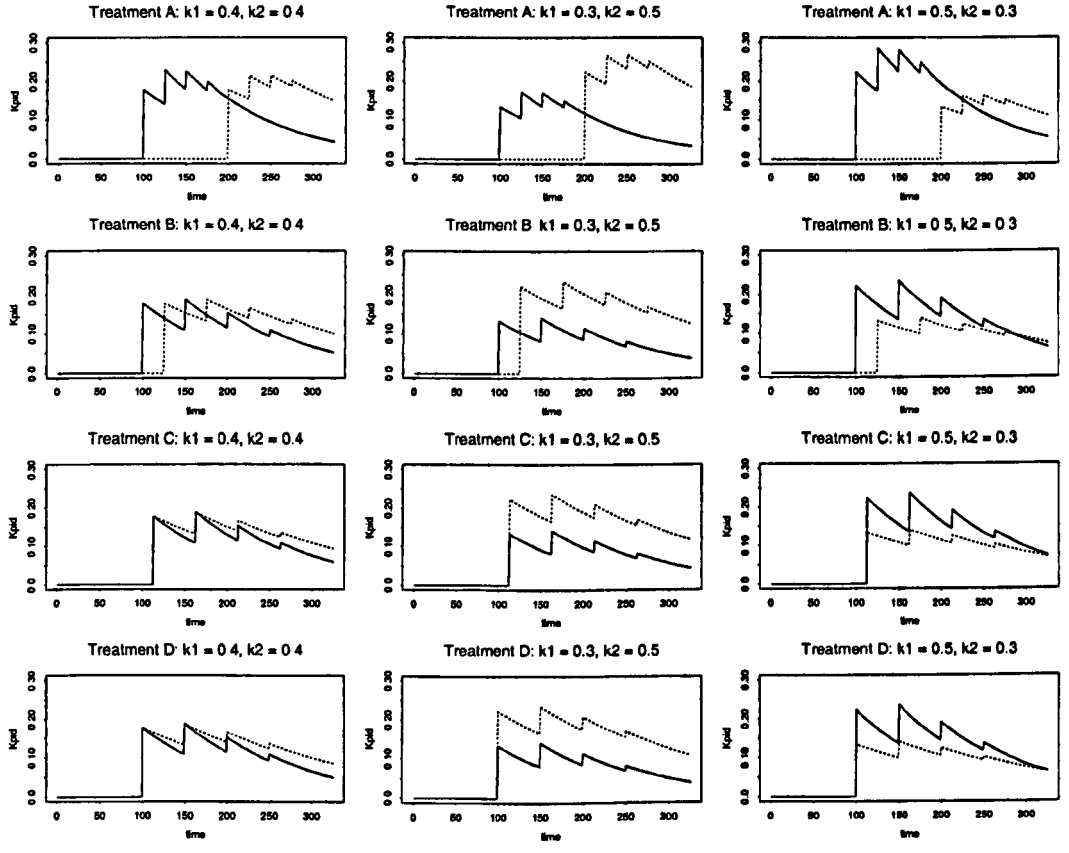
where $k_{pid}(t)$ is given by equation (X) and $C_{e_l}(t)$ needs to be estimated taking into account the pharmacokinetic characteristics of each of the drugs $l = 1, ..., L$.

### Example

We are going to perform a toy experiment in order to see how the model above would behave in a deterministic world, free of measurement error. The evolution of a cell population with the same kinetics will be compared for four different treatment schedules in three different drug potency scenarios.

Let us assume that the population starts with one ancestor at time $t = 1$ and treatment starts at $t = 100$, when the population of leukaemic cells has a critical size. The cell population will be followed up to time $t = 325$.

Figure 6.4: Effect of Drugs 1 and 2, $k_{pid}$, plotted against time for four different treatments (in rows) and three different pair of values of 'maximum effect' of drugs (in columns). Straight lines correspond to the effect of Drug 1 and dashed lines to the effect of Drug 2.



Treatment consists of two different drugs, Drug 1 and Drug 2. The following four doses of each of the drugs must be administered intravenously during treatment:

- Drug 1: $D_1(1) = 40000$, $D_1(2) = 10000$, $D_1(3) = 5000$ and $D_1(4) = 2500$.
- Drug 2: $D_2(1) = 60000$, $D_2(2) = 10000$, $D_2(3) = 5000$ and $D_2(4) = 2500$.

Let us assume that after the first intravenous dose of each of the drugs, the concentration of substances 1 and 2 in the corresponding effect compartment, $C_1(t)$ and $C_2(t)$ satisfy the following kinetic behaviours:

$$\begin{aligned} C_1(t) &= D_1(1)0.5e^{-0.005(t-T_{D_1(1)})} \\ C_2(t) &= D_2(1)0.2e^{-0.003(t-T_{D_2(1)})} \end{aligned}$$

for $t > T_{D_1(1)}$ and $t > T_{D_2(1)}$ and zero otherwise, where $D_1(1)$ and $D_2(1)$ are the initial doses of Drugs 1 and 2 and $T_{D_1(1)}$ and $t > T_{D_1(1)}$ the corresponding dosing times.

Note that these concentrations are understood to correspond to the site of response. The concentration in other body compartments has not been specified. The proportion of the dose that actually reaches the corresponding effect compartment is higher for Drug 1 than for drug 2.

Next, the additive pharmacodynamic model described by equation (6.8) is applied to relate the drug-induced cell death rate to the assumed concentration of drugs in the corresponding effect compartment. It is supposed that for Drug 1, $C_{\text{MAX},1} = 30000$ and that for Drug 2, $C_{\text{MAX},2} = 20000$.

Four different dosing schedules will be considered in this illustration: treatments A, B, C and D. In Treatment A, all four doses of Drug 1 are administered

at times $t = 100, 125, 150$ and 175, followed by the four doses of Drug 2 at times $t = 200, 225, 250$ and 275.

In Treatment $B$ drugs 1 and 2 are administered alternatively, starting with a dose of drug I 1 at time $t = 100$, at 25 time unit intervals. Lastly, Treatments C and D combine both drugs and doses of both substances are given simultaneously at every 50 time unit intervals, starting at time $t = 112.5$ in Treatment C and at time $t = 100$ in Treatment D.

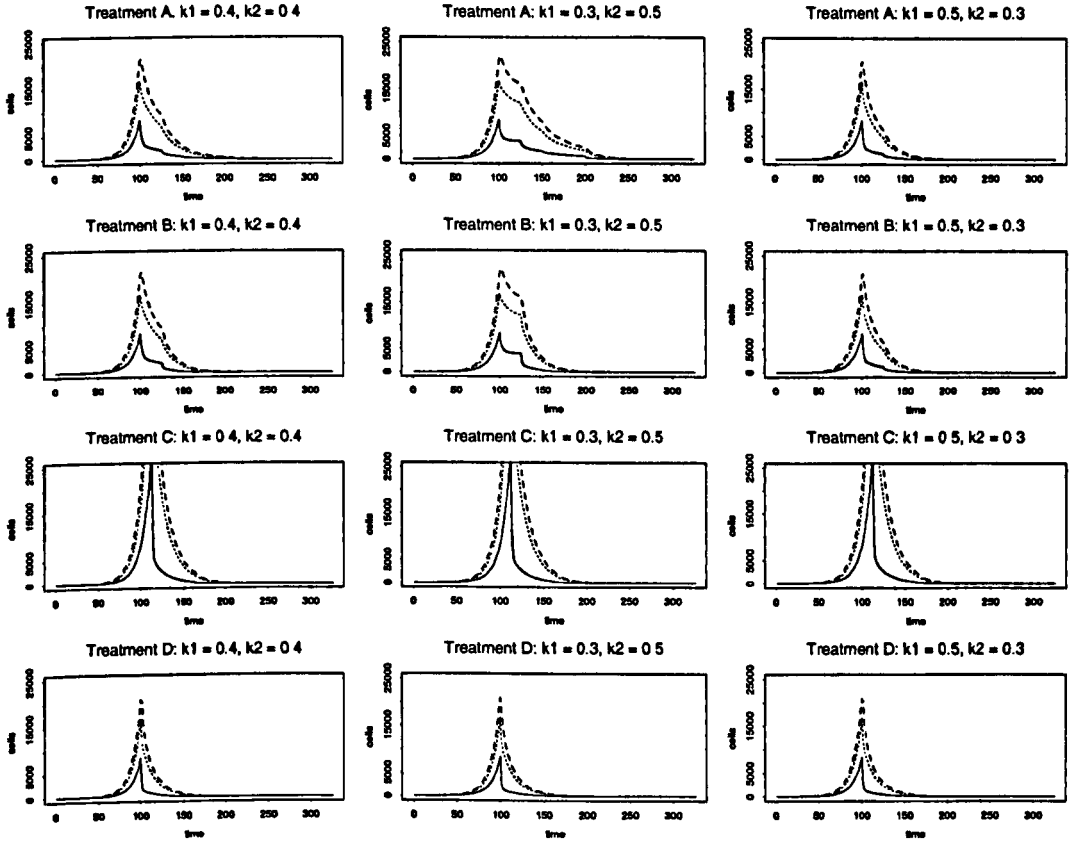
In figure 6.4, the effect $k_{pid}(t)$ satisfying equation (6.8), is plotted against time for all four different treatments for three different pair of values of the maximum effect of the drugs. In the left column the same potential efficiency is assigned to both drugs, in particular, it is assumed that $k_1 = 0.4$ and $k_2 = 0.4$. Higher potential efficiency is given to Drug 2 in the plots in the middle column, fixing $k_1 = 0.3$ and $k_2 = 0.5$. Lastly, in the right column is assumed that $k_1 = 0.5$ and $k_2 = 0.3$. Straight lines correspond to the effect of Drug 1 whereas dashed lines correspond to the effect of Drug 2.

Note that the chosen response variable, $k_{pid}$, due to the assumed pharmacodynamic equation (6.8), is proportional to the drug concentration in the effect compartment. Proportionality with respect to the concentration in the central compartment, which is observable, by no means should be presupposed. The relationship between drug concentration in the central compartment and the effect variable will be characterised by the location of the effect compartment and the kinetics between them, as already discussed.

As mentioned before, the origins of the population are to be found at time $t = 1$ with one ancestor. Before treatment starts, based on the model described in figure 6.3, it is assumed that the evolution of the cell population is governed by the following rate constants:

$$k_{pp} = 0.4, \quad k_{pr} = 0.2, \quad k_{rp} = 0.1, \quad k_{pi} = 0.2, \quad k_d = 0.3.$$

Figure 6.5: Evolution of cell subpopulations under four different treatments (in rows) and three different pair of values of 'maximum effect'(in columns). The rate constants are fixed at $k_{pp} = 0.4$, $k_{pr} = 0.2$, $k_{rp} = 0.1$, $k_{pi} = 0.2$ and $k_d = 0.3$. Straight lines represent the number of reproductive cells, the difference between dotted and straight lines accounts for the number of cells reversibly resting, and the difference between dotted and dashed lines represents the number of irreversibly differentiated cells in the population.



In the plots in figure 6.5 are represented the amount of cells over time in the three subpopulations that constitute the cell population. Straight lines represent the number of reproductive cells. The difference between dotted lines and straight lines accounts for the number of cells that are reversibly resting at each time. Lastly, the difference between dotted and dashed lines represents the number of differentiated cells that will eventually disappear from the organism.

Due to the values assigned to the constant rates, the population of cells increases exponentially prior to treatment. When treatment starts, over 20000 cells constitute the whole population, where approximately half of them are of the reproductive type. As the reproductive subpopulation increases, so does the resting subpopulation, as the fixed constant rate $k_{pr}$ is greater than $k_{rp}$.

When treatment starts, drugs start acting on the reproductive subpopulation, killing cells which otherwise would split. Both the resting and the differentiated subpopulations will eventually decrease accordingly, provided that the total death rate exceeds the splitting rate.

The first conclusion that can be derived from figure 6.5 is that the treatment that best performs in all four scenarios is Treatment D. It should be pointed out that under treatment D not only the total concentration of drugs is higher when the population is bigger, but also that the overall exposure of drugs is greater due to greater total 'areas under the curve'. Superiority of Treatment D could have been predicted since both drugs are administered earlier and, therefore, the total drug exposure is greater.

In Treatment C both drugs are administered simultaneously but treatment starts 25 time units after diagnosis. This time gap can be crucial as the size of the population could reach lethal levels before treatment starts. However, when drug administration begins the number of cells decreases rapidly.

In Treatment A new doses are administered when high amounts of previously administered doses have not yet been eliminated from the body so that drug concentration levels reach high levels but with a shorter exposure length than with other treatments. With Treatment B patients have longer exposure to Drug 2 as its administration starts sooner and the concentration levels of both drugs in the blood are more homogeneous. The performances of Treatments A and B are rather similar. Minor differences arise depending on the efficacy of the drugs, suggesting that the more efficient a drug the better to introduce it sooner in treatment.

Let us imagine that Drug 1 is highly toxic with a high probability of lethal toxicity above certain level. Figure 6.5 shows that under Treatment A, Drug 1 concentration levels are higher than in Treatments B, C and D due to shorter dosing intervals and, therefore, the probability of toxicity higher. Consequently, in this particular setting, the least effective treatment in terms of cell killing rate, would originate the highest probability of lethal toxicity.

Therefore, the model we are considering shows that when setting treatments two intuitive factors need to be considered. First, treatment should start as soon as possible after diagnosis. And, second, the most effective drugs should be administered in early phases of treatment.

The deterministic models that we have built in this toy experiment constitute a good approximation of the evolution of real cell populations in early phases of treatment as it will be shown in chapter 7 for the particular case of leukaemic cell populations.

### 6.2.4 Drug-Resistant Population

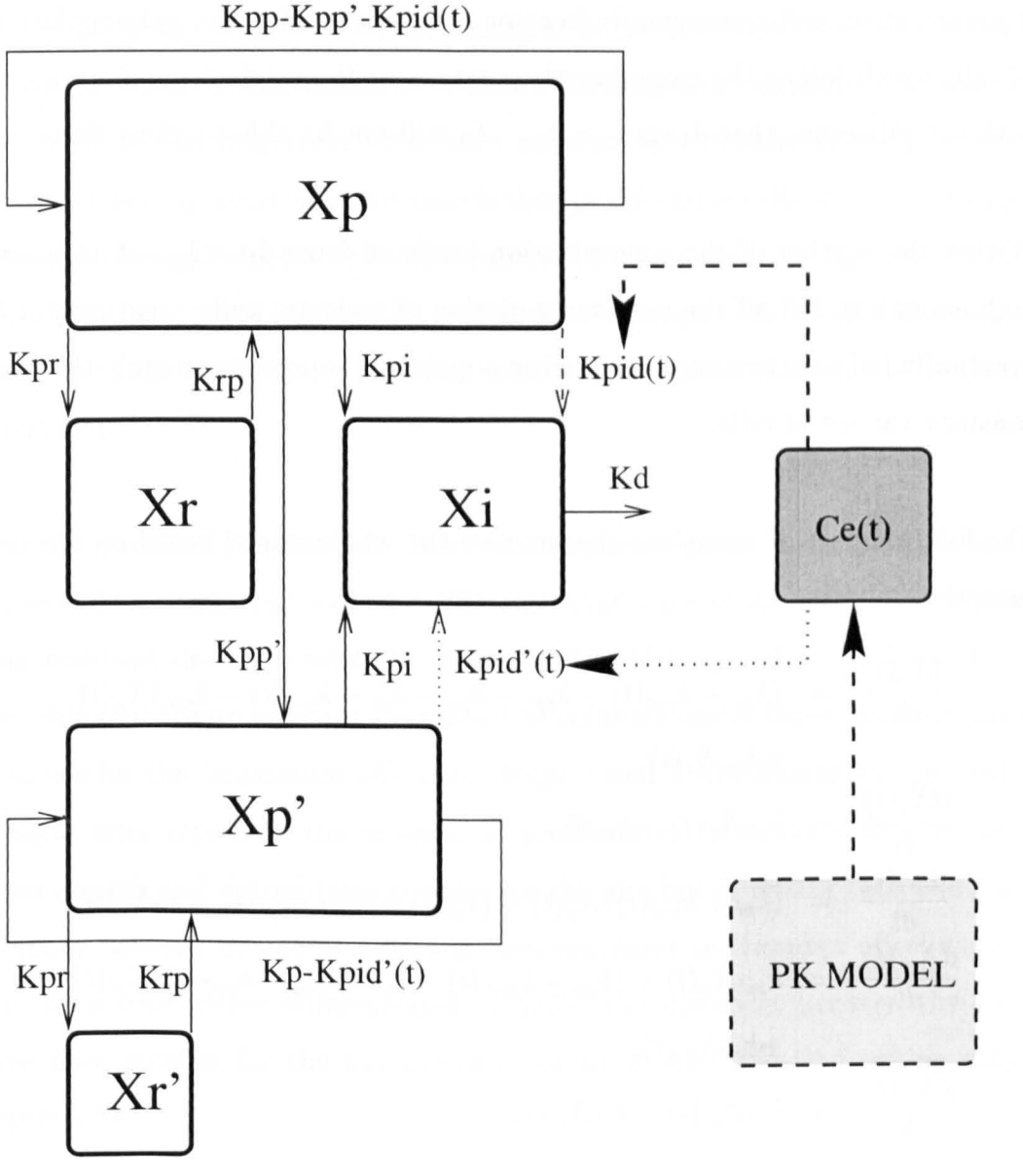
In the model above we have been concerned with the response of subpopulations present in the malignant cell population at the initiation of therapy.

It is well known, however, that the drug response pattern is not necessarily static. New resistant variants may arise due to mutations and other genetic alterations at any time during treatment. The variants, if proliferative, may in time come to comprise new subpopulations within the evolving population and influence therapeutic effectiveness.

In figure 6.6 a model that accommodates drug resistant cells is presented. We start with the model presented in chapter 5 to describe the evolution of cell populations consituted by proliferative, resting and cells that are irreversibly out of the cycle, denoted by $X_p$, $X_r$ and $X_i$ respectively, with the corresponding migration rates. When treatment starts, the pharmacokinetic model will determine the concentration of drugs in the effect compartment, $C_e(t)$. As a result of these concentration levels, a fraction $k_{pid}(t)$ of the proliferative cells that otherwise would split into two daughter cells at rate $k_{pp}$, will leave the cell cycle irreversibly, so that the new splitting rate will be $k_{pp} - k_{pid}(t)$. Nevertheless, a fraction of the original proliferative cells will become resistant to some or all the drugs involved in treatment. For simplicity, in figure 6.6 it is supposed that only one resistant variant arises during treatment, and that they do it at rate $k_{pp'}$. Therefore, the splitting rate of the proliferative cells is now $k_{pp} - k_{pp'} - k_{pid}(t)$.

The number of proliferative resistant cells at time $t$ is denoted by $X_{p'}(t)$ and the corresponding number of resistant cells which decide to rest reversibly by $X_{r'}(t)$. It is assumed that the migration rates within the resistant subpopulation are the same as in the original subpopulation before treatment. The concentration of drugs in the response site, determined by the pharmacokinetic characteristics of each of the drugs, is denoted by $X_e(t)$. Note that the drug-induced death rate of the resistant cells differs from that of the original subpopulation. In fact, $k_{pid'}(t) < k_{pid}(t)$ for all $t$. When cells become resistant to all drugs involved in treatment, $k_{pid'}(t) = 0$. Therefore, the splitting rate of

Figure 6.6: Pharmacokinetic-Pharmacodynamic-Cell-Kinetics Model with a Drug-Resistant Subpopulation. The population before treatment is constituted by three types of cells, proliferative cells, $X_p$, resting cells, $X_r$, and cells which are irreversibly out of the cycle $X_i$, with the corresponding migration rates given by $k_{pr}, k_{rp}, k_{pi}, k_d$ and splitting rate $k_p$. When treatment starts, the pharmacokinetic characteristics of the drugs will determine the concentration of drug in the effect compartment, $C_e(t)$, that will have an effect on the cell population by killing proliferative cells that otherwise would split. In particular, we assume that a fraction $k_{pid}(t)$ of the proliferative cells will not divide during mitosis. In addition, a fraction $k_{pp'}$ is assumed to become resistant to some of the drugs involved in the treatment. These resistant subpopulation will follow the same kinetics as the original subpopulation with the difference that only a fraction $k_{pid'}(t)$ will be killed by treatment, where $k_{pid'}(t) < k_{pid}(t)$.



resistant proliferative cells is given by $k_p - k_{pid'}(t)$ which is going to be smaller than that of the original population.

To make the exposition simpler, let us assume that treatment consists of two groups of drugs, $l_1 = 1, ..., L_1$ and $l_2 = 1, ..., L_2$, such that $L = L_1 + L_2$, where $L$ is the total number of drugs being administered during treatment.

At any time, a proliferative cell from the original population whose predecessor has been exposed to the action of all $L$ drugs, due either to a mutation or to any other genetic alteration, may get resistant to drugs $l_2 = 1, ..., L_2$. This cell, if proliferative, will consequently become the founder of a new subpopulation of cells which follow the same kinetic patterns as the original population but with the difference that drugs $l_2 = 1, ..., L_2$ will not be able to affect them.

Within this setting, if the concentration levels of drugs $l_1 = 1, ..., L_1$ are not high enough to kill off the new supopulation of resistant cells, treatment will eventually fail to extermine the growing population comprising mainly the new resistant variant of cells.

The following set of equations characterise the whole model based on the assumptions above:

$$\begin{aligned}
\frac{dX_p(t)}{dt} &= (k_{pp} - k_{pid}(t) - k_{pp'} - k_{pr} - k_{pi} - k_{pid}(t) - k_{pp'})X_p(t) \\
&\quad +k_{rp}X_r(t), \\
\frac{dX_r(t)}{dt} &= k_{pr}X_p(t) - k_{rp}X_r, \\
\frac{dX_i(t)}{dt} &= (k_{pi} + k_{pid}(t))X_p(t) + (k_{pi}(t) + k_{pid'})X_{p'}(t) - k_dX_i(t), \\
\frac{dX_{p'}(t)}{dt} &= k_{pp'}X_p(t) + (k_{pp} - k_{pid'}(t) - k_{pr} - k_{pi} - k_{pid'}(t))X_{p'}(t) \\
&\quad +k_{rp}X_{r'}(t), \\
\frac{dX_{r'}(t)}{dt} &= k_{pr}X_{p'}(t) - k_{r'}X_{r'}(t),
\end{aligned}$$

where the pharmacodynamic equations are assumed to be:

$$k_{pid}(t) = \sum_{l=1}^{L_1} k_l \frac{C_{e_l}^2(t)}{C_{\mathrm{MAX},l}^2} + \sum_{l=1}^{L_2} k_l \frac{C_{e_l}^2(t)}{C_{\mathrm{MAX},l}^2},$$

$$k_{pid'}(t) = \sum_{l=1}^{L_1} k_l \frac{C_{e_l}^2(t)}{C_{\mathrm{MAX},l}^2},$$

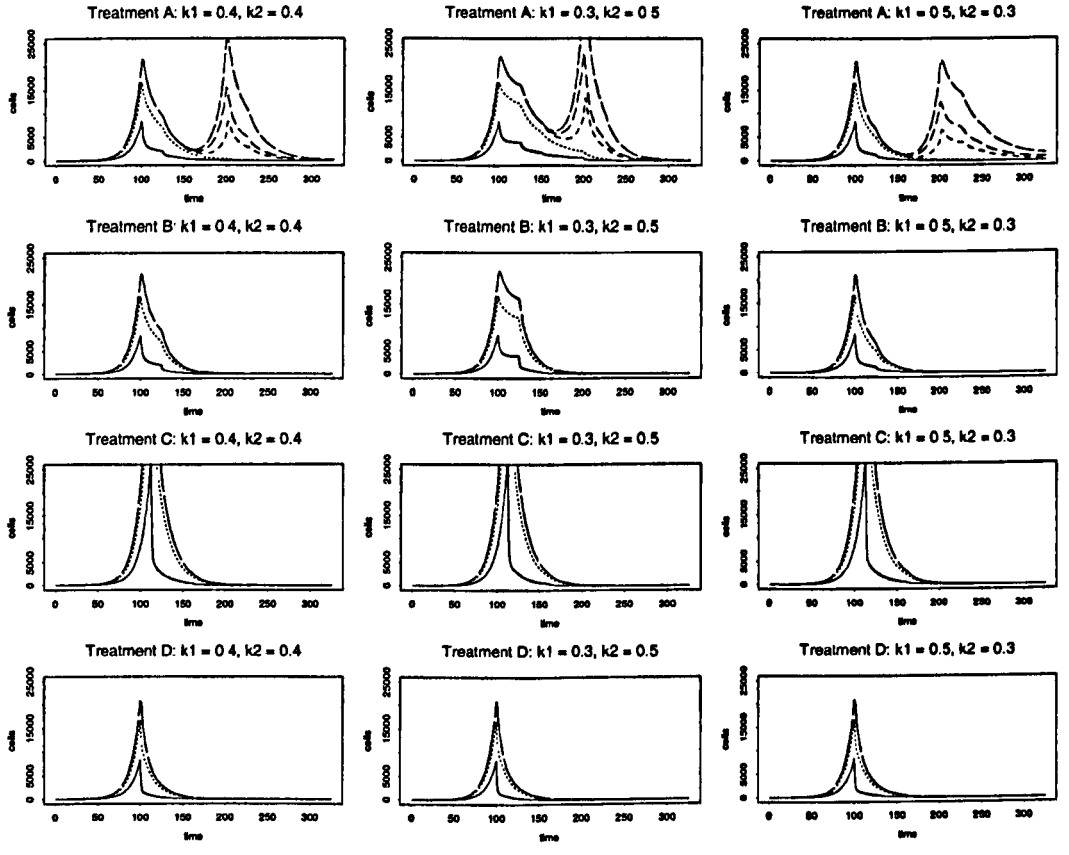
where the notation has already been introduced in the previous section.

### Example

We are going to continue with the example in the previous section by introducing the possibility of proliferative cells becoming drug resistant. Treatment specifications and migration rates have already been defined in the previous section. The only novelty in this case is that proliferative cells which are being exposed to Drugs 1 and 2 will become resistant to Drug 1 at rate $k_{pp'} = 0.00003$. Therefore, Drug 1 will not be able to kill these resistant cells, so that the splitting rate of the proliferative resistant cells will be higher than in the original population.

As in the previous section, four different treatments under three different scenarios will be considered with the difference that some of the cells will become drug resistant during treatment. Figure 6.7 is the equivalent to figure 6.5. Four different treatments are considered in columns under three different pair of values for the 'maximum effect' of Drugs 1 and 2 in columns as specified. Straight lines represent the amount of proliferative cells, the difference between straigh and dotted lines account for the number of resting cells and the diference between dotted and dashed lines represent the number of cells that have been irreversibly differentiated. Finally, the distances between the top three lines account for the number of resistant proliferative and resting cells respectively.

Figure 6.7: Evolution of cell subpopulations under four different treatments (in rows) and three different pair of values of 'maximum effect' (in columns), with a drug resistant subpopulation. The rate constants are fixed at $k_{pp} = 0.4$, $k_{pr} = 0.2$, $k_{rp} = 0.1$, $k_{pi} = 0.2$, $k_d = 0.3$ and $k_{pp'} = 0.00003$. Straight lines represent the number of reproductive cells, the difference between dotted and straight lines accounts for the number of cells reversibly restings, and the difference between dotted and short dashed lines represents the number of irreversibly differentiated cells in the population. The distances between the top three dashed lines account for the number of resistant proliferative cells and the number of resistant resting cells respectively.



In this new setting, it is again Treatment D the one that performs the best, as initial high concentration levels of Drug 2 manage to avoid the growth of the resistant subpopulation. This should not be a surprise as this treatment is the one that introduces both drugs right after diagnosis.

Remember that under Treatment A Drug 2 is not administered until $t = 200$. The resistant population manages to grow rapidly to probably lethal levels when the efficacy of Drug 1 is not high enough (left and center plots).

In all four cases, when Drug 2 is introduced in treatment, the resistant population will decrease rapidly, but this might already be too late as the resistant population might have already reached lethal levels.

As in the previous section under Treatment $C$, which does not administer any drug until $t = 112.5$, cell populations can reach a lethal size before treatment starts.

## 6.3 Population Pharmacodynamic Models

### 6.3.1 Introduction

Recall that in contraposition to pharmacokinetics, the study of what the body does to a drug over time, pharmacodynamics is the study of what the drug does to the body. Substantial variability in effect in a group of similarly treated patients motivates pharmacodynamic investigations. Variability in drug effect may be the result of pharmacokinetic and/or pharmacodynamic factors, and although theoretically of differing biological mechanisms, these factors may not be easily distinguished.

### 6.3.2 Variability in Data

As occurs with pharmacokinetic data, pharmacodynamic data are usually very sparse, both between and within patients, due to different factors. First of all, the effect variable that is being monitored is believed to be a function of the pharmacokinetic processes. These processes, as discussed in earlier chapters, are highly heterogeneous both between and within patients. Even after accounting for most of this variability with an appropriate pharmacokinetic model, the pharmacodynamic model will inheritage the remaining variability.

On the other hand, patients with similar observed drug concentration levels may exhibit substantial differences in the response of their leukaemic cell populations to drug exposure. Interpatient variability that persists after accounting for drug exposure indicates that further statistical modelling with pharmacodynamic factors needs to be done.

On top of that, the response originated by a given drug concentration level may change over time within patients due to, for instance, prior treatments or biological changes in patients' organisms.

The reasons for studying variability in effect are mainly twofold: biologic relevance and clinical applicability. Identification of the biologic factors that cause variability in similarly treated patients has great importance on its own. Directly linked to the biological knowledge are the widespread clinical applications of pharmacodynamic analyses. When pharmacokinetic and/or pharmacodynamic variation leads to an appreciable risk of lethal toxicity at what is assumed to be a safe dose, a modification in administration is warranted.

Pharmacodynamic models are being employed to individualise drug doses such that a safe and effective dose can be given to an individual patient. Ideally, a

pharmacodynamically guided dose will yield improved efficacy while simultaneously decreasing the risk of unacceptable toxicity.

### 6.3.3 Direct Population Pharmacodynamic Models

The population approach to deal with sparse data has been introduced in chapter 3 applied to pharmacokinetic data. As discussed in the previous section, variability in pharmacodynamic data is closely related to that in pharmaokinetic data. Therefore, it seems natural to design a model that embraces both data sets at the same time.

For each patient we have two series of measurements taken over time, one of drug concentration levels observed in blood samples and one of the chosen response variable. Records of the treatment history are usually kept as well. Further individual information that might be relevant in explaining the observed variability is in some cases available.

Following equivalent arguments to those in Chapter 3 for pharmacokinetic models, we are going to build up a model that combines both the pharmacokinetic and the pharmacodynamic processes when no relevant covariate information is available.

A detailed specification of a Pharmacokinetic-Pharmacodynamic model, and a description of alternative methods of estimation of the three stage hierarchical model we briefly present in this section, can be found in Wakefield and Racine-Poon (1995). They apply this methodology to a data set of the drug Recombinant Hirudin.

Let $c_{ij}$ be the $j^{th}$ measured concentration for patient $i$ at time $t_{ij}$, and $y_{ij}$ the reciprocal of the measured effect variable, $i, j = 1, ..., n$. Note that observations of the direct effect of the concentrations are available. There may

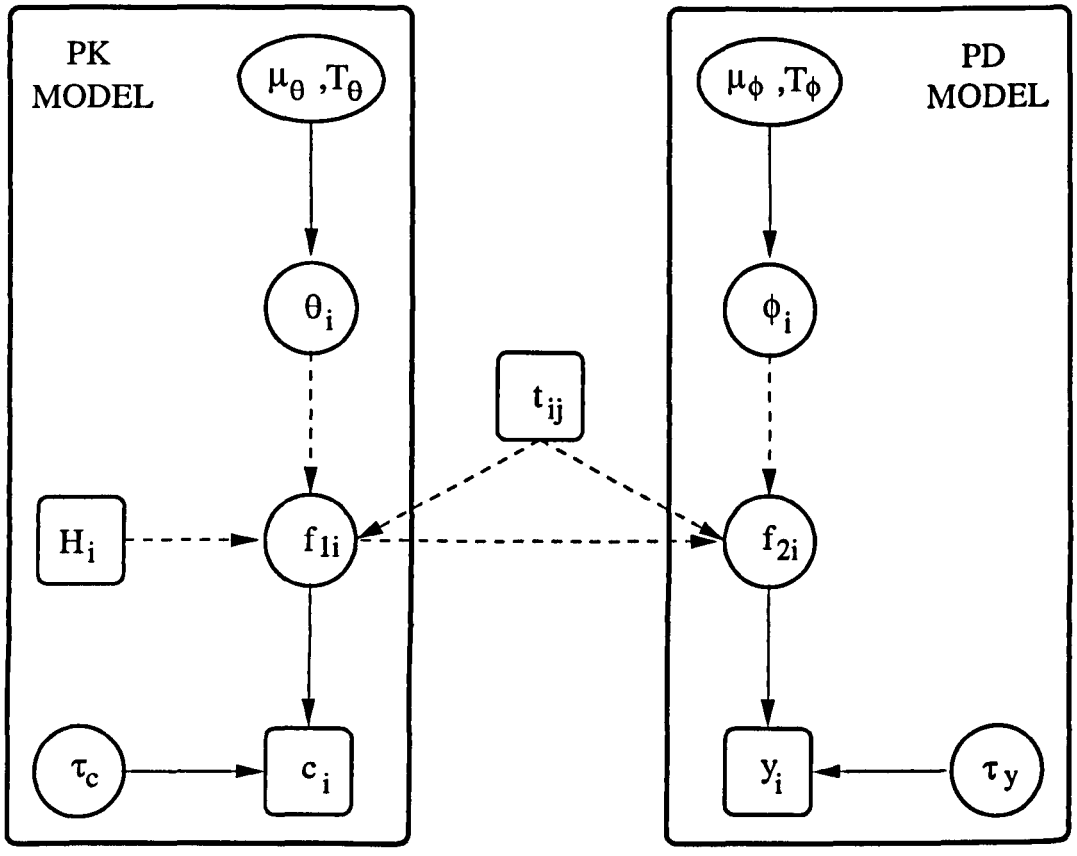



Figure 6.8: DAG representation for the General Pharmacokinetic-Direct-Pharmacodynamic Model. Solid and dashed arrows represent probabilistic and deterministic relationships respectively. Circles represent unknown random quantities whereas fixed or observed quantities are represented by rectangles. Stacked 'plates' represent repetitive structures.

be missing values at some time points, as pharmacokinetic and pharmacodynamic data can be collected at different time points. Let $\theta_i$ and $\phi_i$ be the individual specific vectors of pharmacokinetic and pharmacodynamic parameters respectively. Denoting by $c_i$ and $y_i$ the concentration and effect data vector of individual $i$ and by $\tau_c$ and $\tau_y$ the corresponding intra-individual precision parameters, the likelihood function for the pharmacokinetic and pharmacodynamic parameters of individual $i$ is then given by:

$$\begin{aligned} l(\theta_i, \phi_i, \tau_c, \tau_y) &= \pi(c_i, y_i|\theta_i, \phi_i, \tau_c, \tau_y) && (6.12) \\ &= \pi(c_i|\theta_i, y_i, \phi_i, \tau_c)\pi(y_i|\theta_i, \phi_i, \tau_y)) && (6.13) \\ &= \pi(c_i|\theta_i, \tau_c)\pi(y_i|\theta_i, \phi_i, \tau_y) && (6.14) \end{aligned}$$

where $\pi$ denotes a density function and the last equality holds because the concentration measures depend only on the pharmacokinetic parameters.

At the first stage of the hierarchy, individual observations are described as follows:

$$\ln c_{ij} = \ln f_1(\theta_i, t_{ij}, H_i) + \epsilon_{1ij}, \quad (6.15)$$

$$\ln y_{ij} = \ln f_2(\phi_i, \theta_i, t_{ij}) + \epsilon_{2ij}, \quad (6.16)$$

where $f_1$ and $f_2$ are the corresponding pharmacokinetic and pharmacodynamic models respectively, the treatment history is described by $H_i$, and $\epsilon_1$ and $\epsilon_2$ account for zero mean measurement errors.

Following the population approach, at the second stage, it is assumed that both $\theta_i$ and $\phi_i$ arise from multivariate distributions with mean vectors, say, $\mu_\theta$ and $\mu_\phi$ and precision matrices $\Gamma_\theta$ and $\Gamma_\phi$ respectively.

Finally, the Bayesian approach requires a third stage where prior distributions need to be assigned to the population parameters $\mu_\theta$, $\mu_\phi$, $\Gamma_\theta$, and $\Gamma_\phi$, and to the intra-individual precision parameters $\tau_c$ and $\tau_y$.

Figure 6.8 shows the *Directed Acyclic Graph* (Lauritzen, 1996) which represents this model. DAG representation has already been explained in section 3.3. Priors have not been included in the graph.

Within the Bayesian framework, inference about the posterior distributions of the unknown parameters of the model is targeted. In this case, defining $\theta = (\theta_1, ..., \theta_n)$ and $\phi = (\phi_1, ..., \phi_n)$, the posterior distribution is given by:

$$\begin{aligned}\pi(\theta, \phi, \tau_c, \tau_y, \mu_\theta, \mu_\phi, \Gamma_\theta, \Gamma_\phi) \quad \propto \quad & \prod_{i=1}^{n} \pi(c_i|\theta_i, \tau_c)\pi(y_i|\theta_i, \phi_i, \tau_y) \\ \times \quad & \prod_{i=1}^{n} \pi(\theta_i|\mu_\theta, \Gamma_\sigma)\pi(\phi_i|\mu_\phi, \Gamma_\phi)\pi(\mu_\theta, \mu_\phi, \Gamma_\theta, \Gamma_\phi, \tau_c, \tau_y),\end{aligned}$$

where the notation is the usual.

In Chapter 3 some MCMC techniques which allow sampling from posterior distributions, have been introduced. A Gibbs sampling approach, for instance, would require sampling from the following posterior conditional distributions:

$$\begin{aligned}&\pi(\theta_i|\theta_j, j \neq i, \phi, \tau_c, \tau_y, \mu_\theta, \mu_\phi, \Gamma_\theta, \Gamma_\phi, c, y), \\ &\pi(\phi_i|\phi_j, j \neq i, \theta, \tau_c, \tau_y, \mu_\theta, \mu_\phi, \Gamma_\theta, \Gamma_\phi, c, y), \\ &\pi(\tau_\theta|\theta, \phi, \tau_\phi, \mu_\theta, \mu_\phi, \Gamma_\theta, \Gamma_\phi, c, y), \\ &\pi(\tau_\phi|\theta, \phi, \tau_\theta, \mu_\theta, \mu_\phi, \Gamma_\theta, \Gamma_\phi, c, y), \\ &\pi(\mu_\theta|\theta, \phi, \tau_\theta, \tau_\phi, \mu_\phi, \Gamma_\theta, \Gamma_\phi, c, y), \\ &\pi(\mu_\phi|\theta, \phi, \tau_\theta, \tau_\phi, \mu_\theta, \Gamma_\theta, \Gamma_\phi, c, y), \\ &\pi(\Gamma_\theta|\theta, \phi, \tau_\theta, \tau_\phi, \mu_\theta, \mu_\phi, \Gamma_\phi, c, y), \\ &\pi(\Gamma_\phi|\theta, \phi, \tau_\theta, \tau_\phi, \mu_\theta, \mu_\phi, \Gamma_\theta, c, y),\end{aligned}$$

where $y = (y_1, ..., y_n)$ and $z = (z_1, ..., z_n)$. The prior densities and the particular forms given to the pharmacokinetic and the pharmacodynamic functions will determine the most convenient sampling technique to be used for each of the posterior distributions of the model parameters.

### 6.3.4 Indirect Population Pharmacodynamic Models

In the previous section the population pharmacokinetic/pharmacodynamic methodology has been described when the pharmacological response, $y$, was directed linked to the concentration levels. In this section, a similar hierarchical model will be presented for the case when the response is indirect, that is to say that the response is a result of several processes, only one of which is altered by drugs.

The exposition will be very brief as the methodology will be similar to the one explained in the previous section. The effect variable we will concentrate on will be the effect of drugs on a population of malignant cells. The population-approach will be adopted so that variability of pharmacokinetic, pharmacodynamic and cell population processes will be aknowledged and correspondingly treated.

Let $c_{ij}$ and $z_{ij}$ be the measured concentration and the number of malignant cells respectively for patient $i$ at time $t_{ij}$, where $i, j = 1, ..., n$. Let $f_1$ be the pharmacokinetic equations, with parameters $\theta_i$. The unobservable direct effect of drugs for patient $i$ at time $t_{ij}$ will be denoted by $y_{ij}$. In the case that concerns us, this direct effect would be the function $k_{pid_{ij}}$ introduced in previous sections and, for example, based on equation (6.5), it could be given by:

$$\ln y_{ij} = \ln k_{pid_{ij}} = f_{2_i} + \epsilon_{2_{ij}} = \ln(k_i \frac{c_{ij}^2}{c_{\text{MAX}_i}^2}) + \epsilon_{2_{ij}} \tag{6.17}$$

where we allow for zero mean measurement error with precision parameter $\tau_y$and $f_2$ is the pharmacodynamic model with parameters $\phi_i = (k_i, c_{\text{MAX}_i})$. Finally, let $f_3$ be the function that describes the cell population, with parameters $\delta_i$ which includes the effect variable $y_i = k_{pid_i}$.

Consequently, the first stage of the hierarchical model, where individual ob-

servations are described, is given by:

$$\ln c_{ij} = \ln f_1(\theta_i, t_{ij}, H_i) + \epsilon_{1_{ij}}, \tag{6.18}$$

$$\ln z_{ij} = \ln f_3(\delta_i, \phi_i, \theta_i, \tau_y) + \epsilon_{3_{ij}}, \tag{6.19}$$

where the multiplicative zero mean error $\epsilon_{3_{ij}}$ has precision parameter $\tau_z$.

At the second stage of the model, it is assumed that the individual parameters parameters $\theta_i$, $\phi_i$ and $\delta_i$ are samples from the corresponding population parameters $\theta$, $\phi$ and $\delta$ with mean vectors $\mu_\theta$, $\mu_\phi$ and $\mu_\delta$, and precision matrices $\Gamma_\theta$, $\Gamma_\phi$ and $\Gamma_\delta$.

The model is completed with the third stage of the hierarchy, where prior distributions must be assigned to the population parameters $\mu_\theta$, $\mu_\phi$, $\mu_\delta$, $\Gamma_\theta$, $\Gamma_\phi$ and $\Gamma_\delta$ and to the interindividual precision parameters $\tau_c$, $\tau_y$ and $\tau_z$.

The corresponding *Directed Acyclic Graph* (Lauritzen, 1996) of this three stage hierarchical model is shown in Figure 6.9. Recall section 3.3 where DAG representation has been introduced. Priors have not been included.

The joint posterior distribution of the parameters of the model from where samples should be taken, using the most appropriate MCMC techniques, is given by:

$$\begin{aligned}\pi(\theta, \phi, \delta, \tau_c, \tau_y, \tau_z, \mu_\theta, \mu_\phi\mu_\delta, \Gamma_\theta, \Gamma_\phi, \Gamma_\delta) \propto & \prod_{i=1}^{n} \pi(c_i|\theta_i, \tau_c)\pi(z_i|\delta_i, \phi_i, \theta_i, \tau_y, \tau_z) \\ \times & \prod_{i=1}^{n} \pi(\theta_i|\mu_\theta, \Gamma_\theta)\pi(\phi_i|\mu_\phi, \Gamma_\phi)[\delta_i|\mu_\delta, \Gamma_\delta) \\ \times & \pi(\mu_\theta, \mu_\phi, \mu_\delta, \Gamma_\theta, \Gamma_\phi, \Gamma_\delta, \tau_c, \tau_y, \tau_z),\end{aligned}$$

where $\theta = (\theta_1, ..., \theta_n)$, $\phi = (\phi_1, ..., \phi_n)$ and $\delta = (\delta_1, ..., \delta_n)$. The prior densities and the particular forms given to the pharmacokinetic, pharmacodynamic and cell kinetic functions will determine the most convenient sampling technique to be used to sample from the posterior conditional distributions.

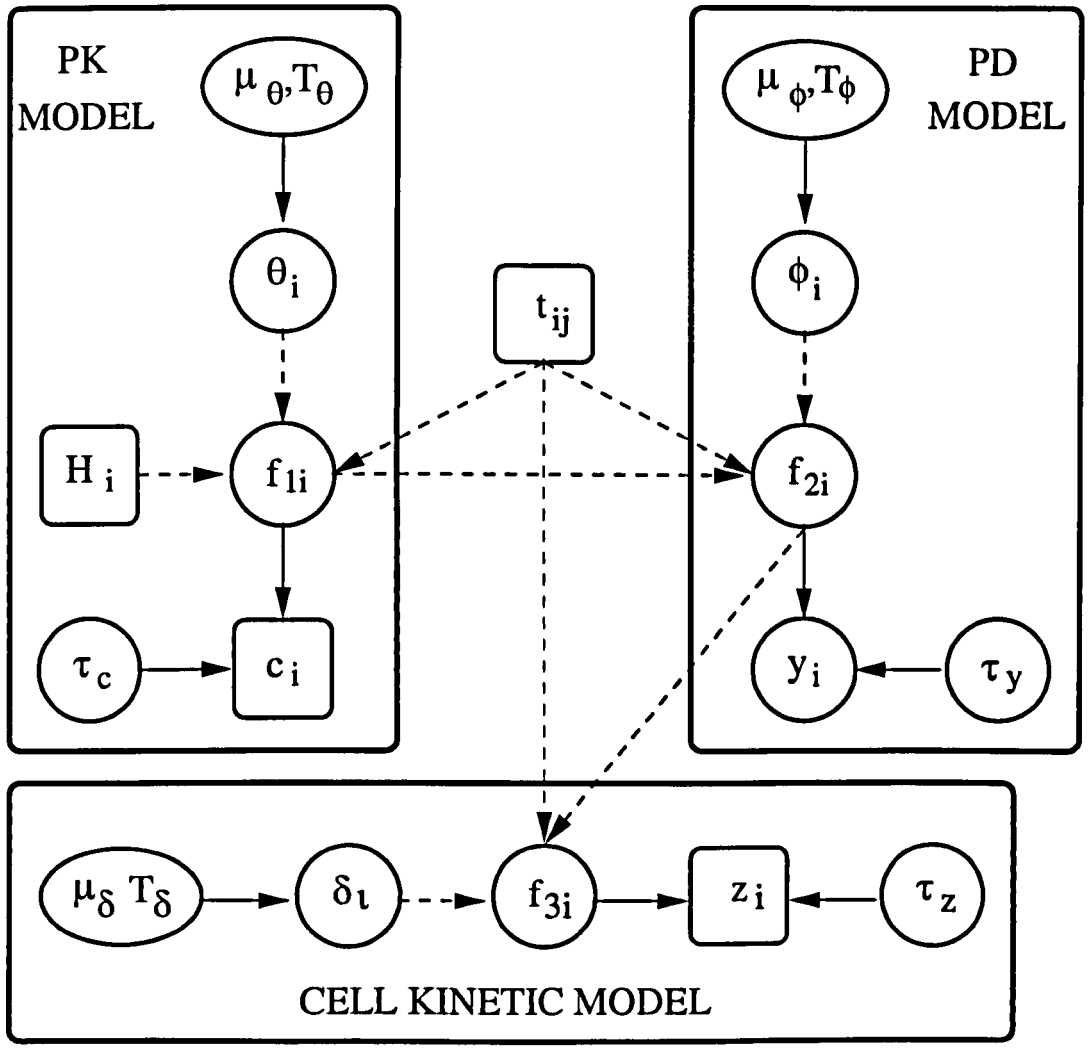



Figure 6.9: DAG representation for the General Pharmacokinetic-Indirect-Pharmacodynamic Model. Solid and dashed arrows represent probabilistic and deterministic relationships respectively. Circles represent unknown random quantities whereas fixed or observed quantities are represented by rectangles. Stacked 'plates' represent repetitive structures.

## 6.4 Conclusions

The main pharmacokinetic processes and models have been described in section 6.2. A differentiation between direct and indirect models has been done in the exposition. Our main interest is to analyse the effect that drugs originate in cell populations and, therefore, considerable time has been spent on indirect models. The relationship between drug concentration levels and the death rate of cells plays a vital role in determining the efficacy of a particular treatment.

Following the approach taken in chapter 3 to model pharmacokinetic processes, a compartmental model has been designed to describe cell populations as a compound of subpopulations with different roles within the cell cycle. The migration rates within and between these compartment have been linked to drug concentration levels in the corresponding effect compartments. Therefore, a threefold model has been built. First, a pharmacokinetic model defining the amount of drug that reaches different parts of the body, including the corresponding effect compartment. Besides, a cell kinetic model describing the evolution of cell populations has been constructed. Thirdly, a pharmacodynamic model has allowed to link drug concentration levels in the effect compartment with the cell kinetic model.

Therefore, a new compartmental Pharmacokinetic-Pharmacodynamic-Cell-Kinetic Model has been constructed. This new approach that we advocate allows to link the evolution of cell populations to particular treatments. We believe that this global view that we have taken can help when designing rational and efficient individual and disease intensity specific treatments.

A fully Bayesian general population three-stage hierarchical model has been specified in section 6.3. This model will be applied to real data in the following chapter.

# Chapter 7

# A Case Study: Leukaemia Treatment

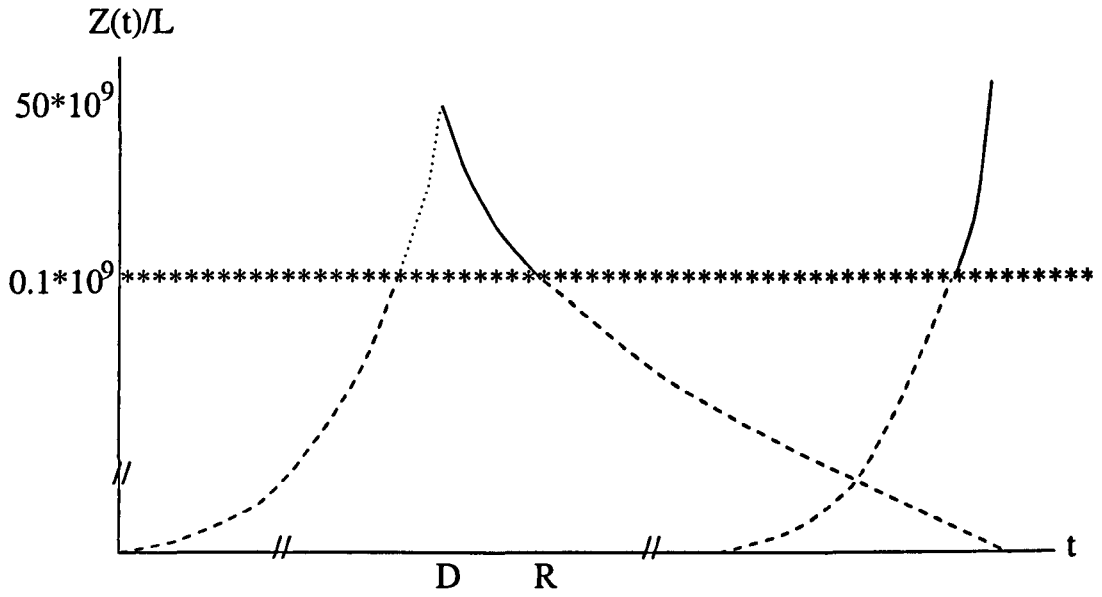
## 7.1 Introduction

In what follows we will study the evolution of a population of leukaemic cells in the circulatory system. Figure 7.1 illustrates a plausible evolution of a leukaemic cell population before and during treatment.

When leukaemia is first diagnosed, the number of malignant cells in the blood ascends to around $50 \times 10^9$ cells per blood liter. Immeditely after diagnosis treatment is initiated. Leukaemia treatment consists of a cocktail of different drugs. Each drug attacks the malignant cells in a particular way. An effective and efficient treatment requires first of all a good understanding of the individual pharmacokinetics of each of the drugs and of the possible intereactions between them.

Once the pharmacokinetic profile of each drug is predicted, the effect of these concentration levels on the leukaemic population should be anticipated in order to administer optimal doses within admissible ranges of toxicity. The current

Figure 7.1: Population of Leukaemic Cells. The disease is diagnosed at time $t = D$, when the population has $50*10^9$ cells per blood liter. Due to treatment, remission will be attained at time $t = R$. Below $0.1 * 10^9$ cells per blood liter, leukaemic cells are unobservable. After remission, a treatment resistant subpopulation will proliferate and eventually reach an observable size.




stage of knowledge is far from being able to set up treatments considering rigorously all these factors.

Virtually all patients enter *remission* within a month. By remission is meant that leukaemic cells are not detectable in blood samples by means of standard techniques. This happens when the number of cells per liter is under $10^9$.

After remission, treatment goes on until, hopefully, no leukaemic cells are left in the organism. The main difficulty lies in the fact that effectiveness is not observable and, consequently, a new population of leukaemic cells, ressistant to that particular treatment, could start forming at any time with no possible detection and, therefore, fatal consequences.

## 7.2 Data

The data set we are concerned with is of the form shown in figure 7.2. This data set was provided by Dr. Stephen Lowis from the Bristol Royal Hospital for Sick Children. For each of the 15 patients in the study, individual treatment details and the number of leukaemic cells in the blood are collected at certain time points. The first column in figure 7.2 gives the date. The second column accounts the number of leukaemic cells per blood liter, in $10^9$ units. The remaining columns describe the administered doses of each of the five drugs that constitute treatment. Individual doses are calculated according to the body mass of the patient, which is not recorded. Dose times and frequency will vary between patients.

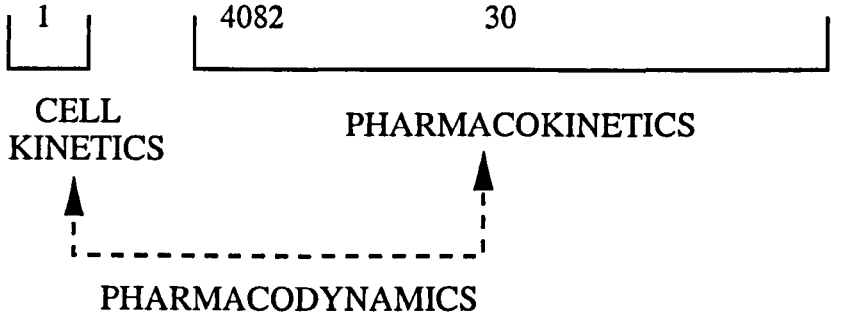
In this data set there are three main interrelated processes that should in principle be modelled in order to be able to estimate and predict the fate of the population of leukaemic cells given the treatment. It constitutes what in the previous chapter has been called as a pharmacokinetic-indirect-pharmacodynamic model.

First of all the concentration of each of the drugs within the body must be studied. This requires a pharmacokinetic analysis for each of the substances. As treatment consists of a cocktail of five different drugs, the pharmacokinetic analysis should account for possible interactions between drugs.

As discussed in previous chapters, given that the pharmacokinetic profiles generally vary between and within individuals, a population pharmacokinetic model could be designed in order to estimate the amount of drug that actually reaches the circulatory system over time for each of the patients. Note that as leukaemic cells are located in the circulatory system, the site of response will be proportional to the concentration in the central compartment.

Figure 7.2: Typical available data set for each of the 15 patients in the study: date, number of leukaemic cells per blood liter, in 10 units, and the doses of each of the five drugs.

| Date | Cells | Asp | VCR | Pred | Dauno | MTX |
|---|---|---|---|---|---|---|
| 24-04 | 6.8 | | 2.5 | 30 | | 12.5 |
| 25-04 | 3.9 | | | 30 | | |
| 26-04 | 3.8 | 4082 | | 30 | | |
| 27-04 | | | | 30 | | |
| 28-04 | 2.1 | 4082 | | 30 | | |
| 29-04 | 1.8 | | | 30 | | |
| 30-04 | 1 | 4082 | | 30 | | |




The first difficulty that arises from the data set is that drug concentrations have not been reported. Therefore, with no observations of blood concentrations for the drugs, assumptions regarding the pharmacokinetic parameters need to be done.

Once the evolution of the concentration of different substances within the circulatory system is available, the effect that these substances originate in different organ and tissues need to be analysed. The effect variable we are interested in is the death rate of leukaemic cells. Therefore, a pharmacodynamic model that links the concentration of the drug cocktail with the evolution of the death rate will be required.

In order to establish a relationship between drugs and cell death, we need to understand how the cell population behaved before treatment started. Therefore, leukaemic cell kinetics need to be studied.

## 7.3 The Structural Model

### 7.3.1 The Pharmacokinetic Model

It has already been mentioned that drug concentrations have not been reported in the data set. Therefore, a pharmacokinetic analysis of the substances cannot be performed.

This lack of vital information will be overcome by means of assuming a one-compartment pharmacokinetic behaviour for the concentration at the site of effect and arbitrarily fixing the values of the corresponding parameters.

However, the aim of this study is not to carry out a pharmacokinetic analysis of the substances that constitute the treatment, but to try to associate qualitatively the rate of death of leukaemic cells with particular treatment schemes. Any quantitative link will rely upon the assumed concentration levels which will most probably differ substantially from the real concentraion levels.

The chosen pharmacokinetic model for all the drugs is a one-compartment model (see section 2.4.2) and we assume that the concentration levels for each of the drugs (Asp, VCR, Pred, Dauno and MTX) are given by:

$$\begin{aligned} C_{ij}^{\text{Asp}} &= d^{\text{Asp}} 0.4 e^{-0.5 t_{ij}} \\ C_{ij}^{\text{VCR}} &= d^{\text{VCR}} 0.2 e^{-0.5 t_{ij}} \\ C_{ij}^{\text{Pred}} &= d^{\text{Pred}} 0.4 e^{-0.5 t_{ij}} \\ C_{ij}^{\text{Dauno}} &= d^{\text{Dauno}} 0.1 e^{-0.5 t_{ij}} \\ C_{ij}^{\text{MTX}} &= d^{\text{MTX}} 0.2 e^{-0.5 t_{ij}} \end{aligned} \tag{7.1}$$

where $C$ is the drug concentration found in the blood, $d$ is the administered dose, $t$ is the time point, subscript $i = 1, ...15$ identifies the patient, $j = 1, ..., 14$ the observation number and the superscript refers to the corresponding drug name.

Figure 7.3: Assumed drug concentration levels for patients 1-8 following a single intravenous administration of each of the administered drugs.

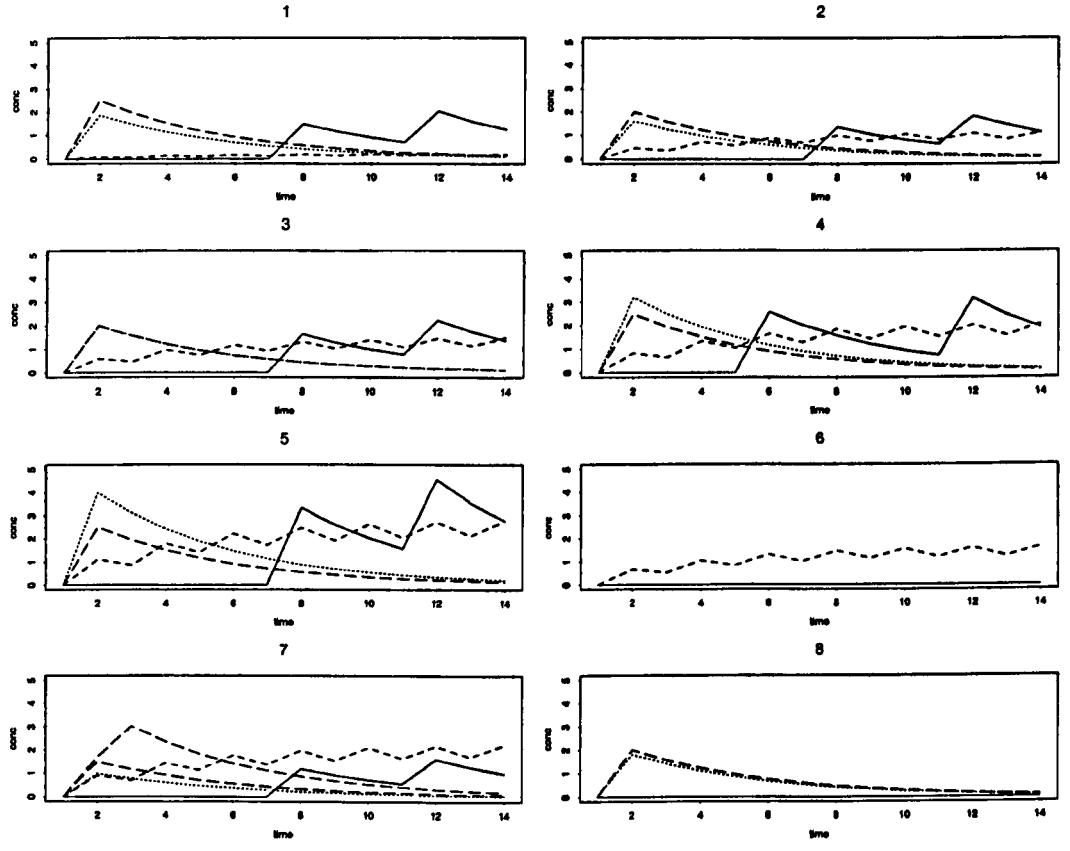
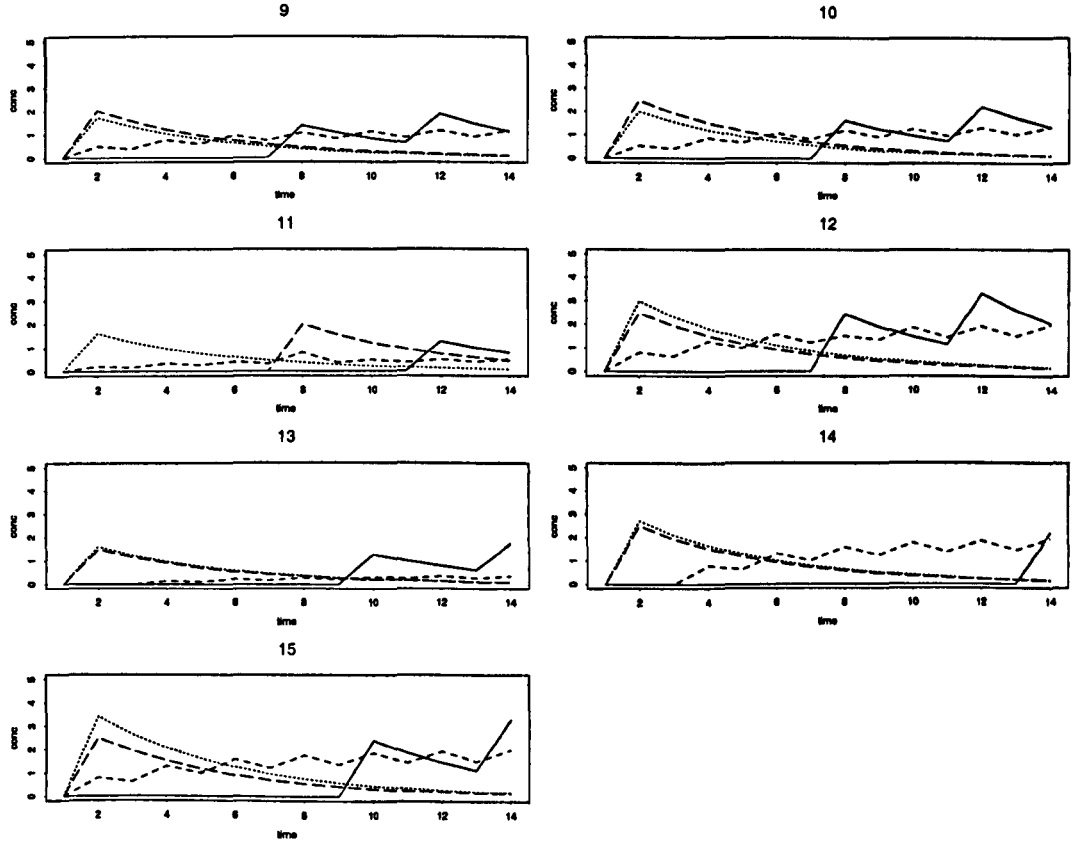


Figure 7.4: Assumed drug concentration levels for patients 9-15 following a single intravenous administration of each of the administered drugs.



There are only observations for seven time points in the data set, one per day. We have decided to add a second daily observation time in order to accommodate possible discrepancies in blood sample and dosing times. Therefore, we assume that blood samples to count the number of malignant cells are taken in the morning and that the administered drugs reach the circulatory system in the afternoon, due to either later dosing times or possible time delays in the pharmacokinetic processes.

The concentration levels derived from these particular exponential equations are presented in figures 7.3 and 7.4. The concentration profiles differ substantially between patients due to dosing differences. Patients 6 and 8 are exposed to particularly low concentrations.

### 7.3.2 The Cell Kinetic Model

As it has been explained in chapter 5, cells live for certain random times and after dying are replaced either by two newborn daughter cells or by none. Different cell kinetic models were presented there.

The theory of Branching Processes presented in Section 5.6 presents the advantage that, involving a relatively moderate number of parameters, it accommodates general life length distributions. However, it should be kept in mind that all results presented there were asymptotic results. In addition, the reproduction mean was supposed to be constant over time, whereas our goal in this section is to establish a relationship between the reproduction mean and the concentration of drug. Changes in concentration should imply changes in the reproduction mean, and consequently in the evolution of the cell population.

Let us denote the $j^{th}$ observed number of cells for patient $i$ by $Z_{ij}$, and the time at which the corresponding sample was collected by $t_{ij}$. We set $t_{i1} = 0$ and

take half-day as time unit [1]. The model we propose to explain the evolution of the leukaemic cell population before treatment starts is the following:

$$E(Z_{ij}) = K_i e^{\bar{\alpha}_i t_{ij}} \tag{7.2}$$

for all $i = 1, 2, ..., 15$ and $j = 1, 2, ..., 14$ and where the intercept $K_i$ and the slope $\bar{\alpha}_i$ are parameters to be determined.

It is believed that the observed number of malignant cells are subject to some sort of measurement error. Since assay techniques are frequently observed to produce measurements whose precision decreases with increasing number of cells, we take logarithms in equation (7.2) and add a zero mean error:

$$\ln Z_{ij} = z_{ij} = K_i + \bar{\alpha}_i t_{ij} + \epsilon_{ij}. \tag{7.3}$$

Note that as we have set $t_{i1} = 0$, $z_{i1} = K_i + \epsilon_{i1}$ for all $i = 1, ..., 14$.

### 7.3.3 The Pharmacodynamic Model

In order to relate drug doses with the cell population, equation (7.3) should depend on drug concentration levels.

We believe that both parameters $K$ and $\bar{\alpha}$ should vary according to concentration levels. Nevertheless, given that at time $t = 0$ treatment has not started yet and that the available number of observations is limitated, our first attempt will be to keep the intercept constant during treatment. Consequently, we will model the slope parameter $\bar{\alpha}$ as a function of drug concentration levels. As these change during treatment, so will do $\bar{\alpha}$.

Let us assume that

$$\bar{\alpha}_{ij} = \alpha_i (A - \sum_l C_{ij}^l), \tag{7.4}$$

[1] Note that for even $j$ blood samples were not collected, so that the number of malignant cells is not available.

where $l$ = Asp, VCR, Pred, Dauno, MTX, and $A$ is the quantity that will determine the sign of the slope $\bar{\alpha}$. In order to incorporate some degree of flexibility to the model, we add a zero mean error in equation (7.4) so that,

$$\bar{\alpha}_{ij} = \alpha_i(A - \sum_l C_{ij}^l) + v_{ij}. \tag{7.5}$$

## 7.4 The Probability Model

In figure 7.5 the population cell-kinetic-pharmacodynamic model we are going to construct is represented as a *Directed Acyclic Graph* (Lauritzen, 1996). DAG representation has been explained in section 3.3. As it will be shown, a three-stage hierarchical model captures all the structural relationships within the model.

The first stage of the hierarchical model is constituted by:

$$z_{ij} = K_i + \bar{\alpha}_{ij} t_{ij} + \epsilon_{ij}, \tag{7.6}$$

$$\bar{\alpha}_{ij} = \alpha_i(A - \textstyle\sum_l C_{ij}^l) + v_{ij}, \tag{7.7}$$

where the notation has already been introduced.

The form of the probability distribution of each $z_{ij}$ given $K_i$, $\bar{\alpha}_{ij}$ and $\tau_z$, the inverse of the residual error, $\epsilon_{ij}$, variance, needs to be specified. We assume that the residual errors $\epsilon_{ij}$ are normally distributed independent errors, so that:

$$[z_{ij}|K_i, \bar{\alpha}_{ij}, \tau_z] = N(K_i + \bar{\alpha}_{ij} t_{ij}, \tau_z^{-1}), \tag{7.8}$$

where as in chapter 4, we follow the notation of Gelfand and Smith (1990) for defining probability densities, i.e., joint, conditional and marginal densities of the random variables $x$ and $y$ are denoted by $[x, y]$, $[x|y]$ and $[x]$,$[y]$ respectively. A normal distribution with mean $a$ and precision parameter $b$ is as usual denoted by $N(a, b)$.

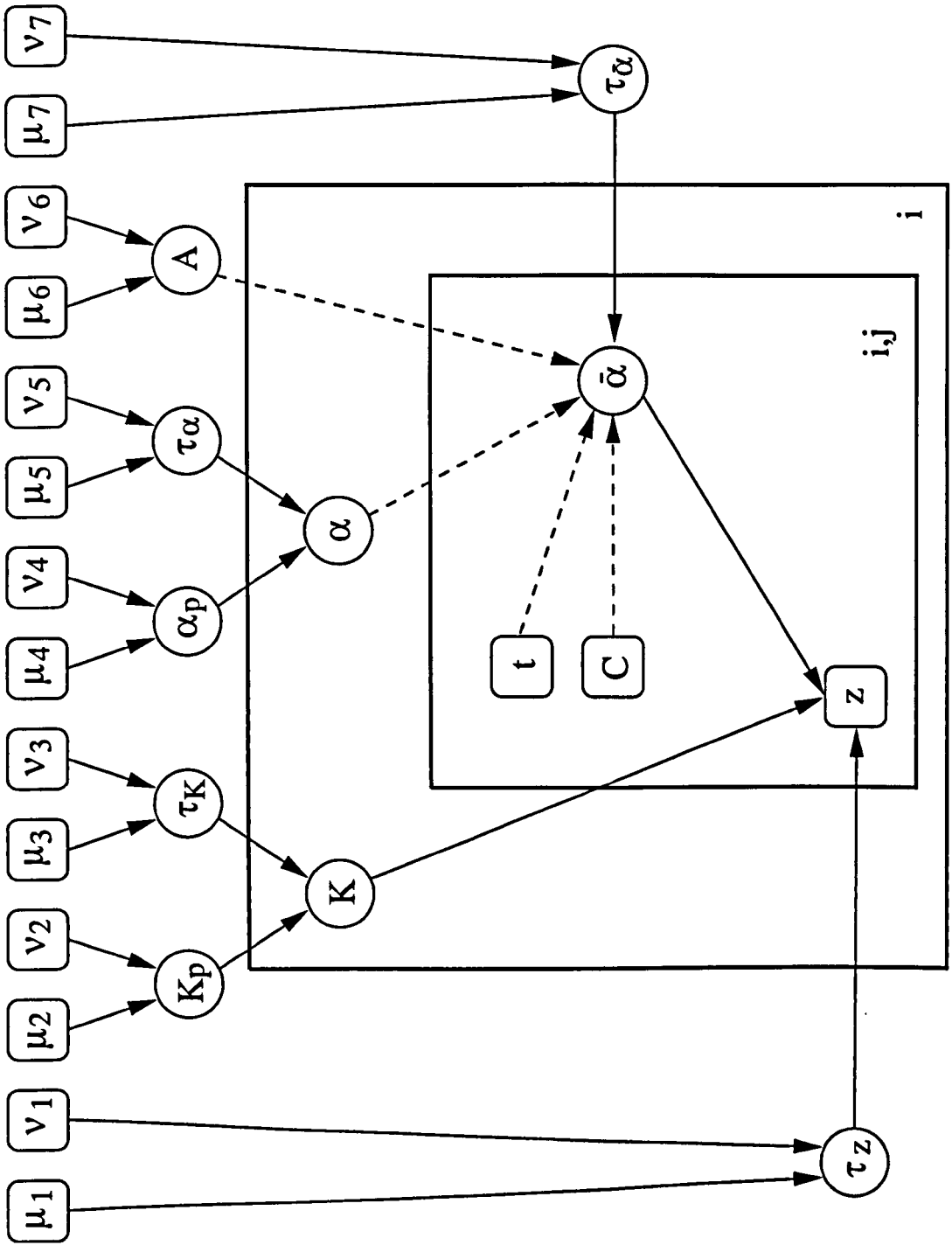



Figure 7.5: DAG representation for the Cyclosporin Model. Due to the lack of relevant covariate information about the patients, the structure of the model is more simple than in the general model. Solid and dashed arrows represent probabilistic and deterministic relationships respectively. Circles represent unknown random quantities whereas fixed or observed quantities are represented by rectangles. Stack 'plates' represent repetitive structures.

Equivalently, assuming that the errors $v_{ij}$ follow a normal distribution with zero mean and precision parameter $\tau_{\bar{\alpha}}$, the distribution of each of the slopes can be written as:

$$[\bar{\alpha}_{ij}|\alpha_i, A, C_{ij}^l, \tau_{\bar{\alpha}}] = N(\alpha_i(A - \sum_l C_{ij}^l), \tau_{\bar{\alpha}}^{-1}). \tag{7.9}$$

From the conditional independence assumption, denoting by $z$ the totality of the observed data, by $K = (K_1, ..., K_{15})$, by $\bar{\alpha}$ all possible slopes and by $\alpha = (\alpha_1, ..., \alpha_{15})$, it follows that,

$$[z|K, \bar{\alpha}, \tau_y] = \prod_{i=1}^{15} \prod_{j=1}^{14} N(K_i + \bar{\alpha}_{ij} t_{ij}, \tau_z),$$
$$[\bar{\alpha}|\alpha, A, \tau_{\bar{\alpha}}] = \prod_{i=1}^{15} \prod_{j=1}^{14} N(\alpha_i(A - \textstyle\sum_l C_{ij}^l), \tau_{\bar{\alpha}}).$$

At the second stage of the probabilistic hierarchical model, assumptions regarding the distribution of the parameters $K = (K_1, ..., K_{15})$ and $\alpha = (\alpha_1, ..., \alpha_{15})$ are specified. We assume, following the population approach, that $\alpha_i$ and $K_i$ are qualitatively the same for all individuals but quantitatively allowed to vary according to the following distributions:

$$[\alpha_i|\alpha_p, \tau_\alpha] = N(\alpha_p, \tau_\alpha^{-1}),$$
$$[K_i|K_p, \tau_K] = N(K_p, \tau_K^{-1}),$$

for $i = 1, ..., 15$, where $\alpha_p$ and $K_p$ are the corresponding population means, and $\tau_\alpha$ and $\tau_K$ the population precision parameters.

To complete the specification of the model, in the third stage of the hierarchy prior densities must be assigned to the parameters without *parents*:

$$[\tau_z] = \text{Ga}(\mu_1, v_1) \quad [K_p] = N(\mu_2, v_2) \quad [\tau_K] = \text{Ga}(\mu_3, v_3) \quad [\alpha_p] = N(\mu_4, v_4)$$
$$[\tau_\alpha] = \text{Ga}(\mu_5, v_5) \quad [A] = N(\mu_6, v_6) \quad [\tau_{\bar{\alpha}}] = \text{Ga}(\mu_7, v_7)$$

where the notation is the conventional used in prior chapters.

Lastly, as discussed in chapters 3 and 4, we need to assign values to the hyperparameters of the model, *i.e.* $\mu_r, v_r$, $r = 1, ..., 7$. As we do not have any

prior information available in our study we introduce very vague prior distributions. The distributions we choose are $N(0, 0.0001)$ and $Ga(0.001, 0.001)$. These distributions are proper but very uninformative.

## 7.5 Estimation

In the former section distributional assumptions of the three-stage hierarchical model have been done. Applying Bayes' theorem, the joint posterior distribution will be proportional to the product of the likelihood function and the prior densities.

We have made use of the software **WINBUGS** (94) to obtain samples from the estimated posterior distributions of the model parameters conditional on the data. To sample from the posterior distributions of the precision parameters the Metropolis-Hastings (Metropolis *et al.* 1953, Metropolis and Hastings, 1970) algorithm is employed, whereas to get samples from the remainder parameters' posterior distributions the Gibbs sampler (Geman and Geman, 1984) is used.

In **WINBUGS** initial values for all the chains are specified. Several different runs starting with different sets of initial values have been performed with virtually identical results. After several attempts, the first 10.000 iterations of the chains were discarded as 'burn-in', in order to, hopefully, guarantee that the stationary distribution is reached. The chains have then been run for additional 20.000 iterations. The computation took around 5 minutes.

In tables 7.1 and 7.2, the estimated posterior marginal distributions of the population and individual slopes and intercepts are summarised. The first two columns give the expected value and the standard deviations of the posterior distributions. The *MC errors* in the fourth column, is the standard error of the

Table 7.1: Summary: Marginal Posterior Distributions for Population, $\alpha_p$, and Individual Slopes, $\alpha_i$, $i = 1, ..., 15$.

| Parameter | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|
| $\alpha_p$ | 0.08069 | 0.04552 | 0.0007694 | -0.006851 | 0.07976 | 0.1747 |
| $\alpha_1$ | 0.1427 | 0.05236 | 0.001192 | 0.0492 | 0.1399 | 0.2535 |
| $\alpha_2$ | 0.0731 | 0.04621 | 0.0009793 | -0.01225 | 0.07049 | 0.172 |
| $\alpha_3$ | 0.04069 | 0.03174 | 0.000554 | -0.01869 | 0.03976 | 0.1056 |
| $\alpha_4$ | 0.05787 | 0.02524 | 0.0006497 | 0.0007127 | 0.04687 | 0.1007 |
| $\alpha_5$ | 0.02095 | 0.04098 | 0.001731 | -0.05967 | 0.02115 | 0.1013 |
| $\alpha_6$ | 0.2176 | 0.1204 | 0.002723 | 0.03444 | 0.1997 | 0.4979 |
| $\alpha_7$ | 0.09555 | 0.02632 | 0.0006203 | 0.04619 | 0.09488 | 0.1502 |
| $\alpha_8$ | -0.02315 | 0.08432 | 0.001968 | -0.1943 | -0.0203 | 0.1332 |
| $\alpha_9$ | 0.1441 | 0.05177 | 0.001169 | 0.04859 | 0.1414 | 0.2524 |
| $\alpha_{10}$ | 0.03538 | 0.03063 | 0.000595 | -0.02368 | 0.03446 | 0.09787 |
| $\alpha_{11}$ | 0.1831 | 0.0576 | 0.001018 | 0.07971 | 0.1799 | 0.3064 |
| $\alpha_{12}$ | 0.04689 | 0.02562 | 0.0004419 | -0.002793 | 0.04644 | 0.09862 |
| $\alpha_{13}$ | 0.05883 | 0.07494 | 0.001593 | -0.07761 | 0.05351 | 0.2215 |
| $\alpha_{14}$ | 0.109 | 0.09183 | 0.005002 | -0.06178 | 0.1042 | 0.3017 |
| $\alpha_{15}$ | 0.02 | 0.03053 | 0.0006257 | -0.03829 | 0.01924 | 0.08264 |

estimate of the expected value in the second column; it is estimated by considering both the sample size and the degree of autocorrelation in the sample. The last three columns give the quantile and the median of the corresponding distributions.

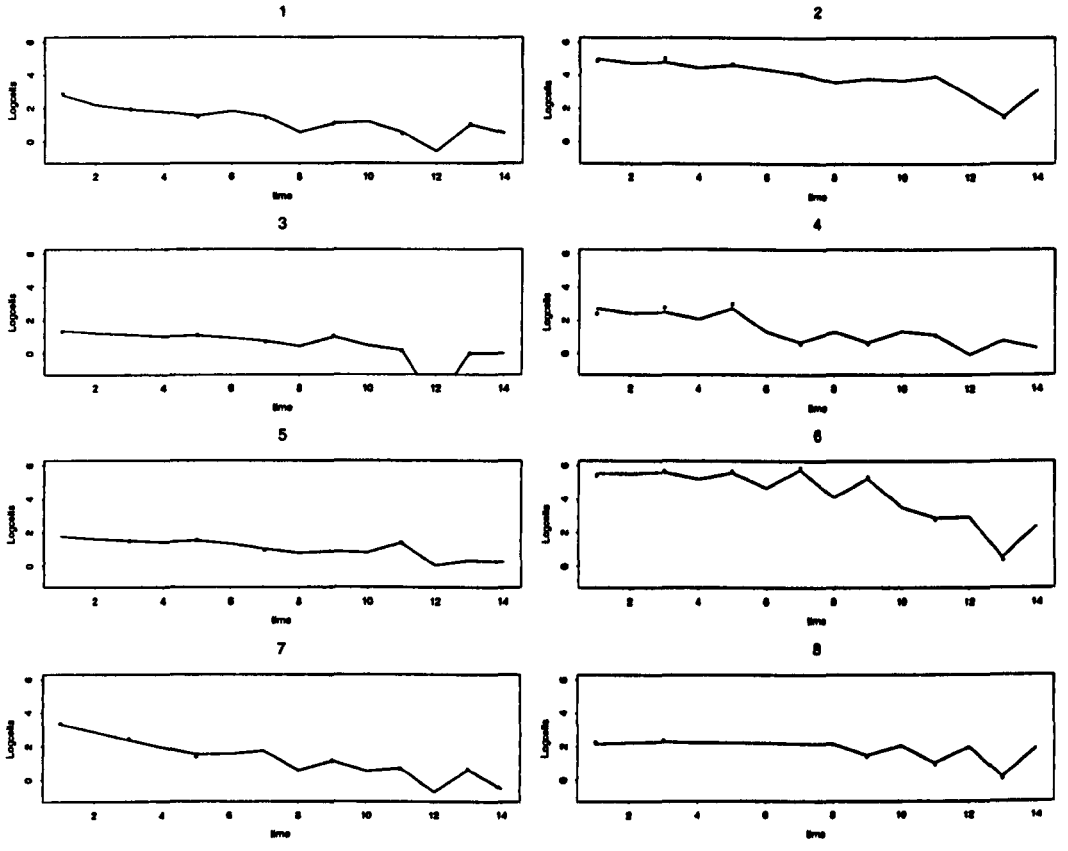
The values of the standard deviations and the quantiles of the posterior distributions in tables 7.1 and 7.2 suggest that these distributions are reasonably symmetric and concentrated around their mean values, as expected. Consequently, and taking in account the small *MC errors* in all cases, we can conclude that the expected values of the posterior distributions constitute safe point estimators of the corresponding real values.

In figure 7.8 we have included the estimated posterior distribution for the population intercept, $K_p$. As it can be seen, the behaviour of this function is

Table 7.2: Summary: Marginal Posterior Distributions for Population, $K_p$, and Individual Intercepts, $K_i$, $i = 1, ..., 15$.

| Parameter | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|
| $K_p$ | 2.438 | 0.4454 | 0.005464 | 1.545 | 2.442 | 3.312 |
| $K_1$ | 2.771 | 0.2653 | 0.006477 | 2.221 | 2.78 | 3.273 |
| $K_2$ | 4.993 | 0.2657 | 0.005365 | 4.468 | 4.989 | 5.537 |
| $K_3$ | 1.37 | 0.2715 | 0.006133 | 0.8412 | 1.365 | 1.921 |
| $K_4$ | 2.728 | 0.2822 | 0.007591 | 2.208 | 2.713 | 3.332 |
| $K_5$ | 1.747 | 0.6477 | 0.02867 | 0.4836 | 1.758 | 3.023 |
| $K_6$ | 5.552 | 0.2373 | 0.005858 | 5.092 | 5.551 | 6.03 |
| $K_7$ | 3.329 | 0.2728 | 0.006354 | 2.756 | 3.336 | 3.854 |
| $K_8$ | 2.149 | 0.2675 | 0.005967 | 1.614 | 2.115 | 2.668 |
| $K_9$ | 3.33 | 0.2779 | 0.00625 | 2.753 | 3.338 | 3.856 |
| $K_{10}$ | 1.116 | 0.2504 | 0.005757 | 0.6214 | 1.115 | 1.619 |
| $K_{11}$ | 3.699 | 0.306 | 0.009471 | 3.032 | 3.723 | 4.232 |
| $K_{12}$ | 1.237 | 0.254 | 0.005058 | 0.7311 | 1.238 | 1.737 |
| $K_{13}$ | 0.9829 | 0.2537 | 0.006055 | 0.4818 | 0.9824 | 1.488 |
| $K_{14}$ | 0.8792 | 0.817 | 0.04519 | -0.7423 | 0.8691 | 2.462 |
| $K_{15}$ | 0.7148 | 0.2788 | 0.005783 | 0.1805 | 0.7041 | 1.293 |

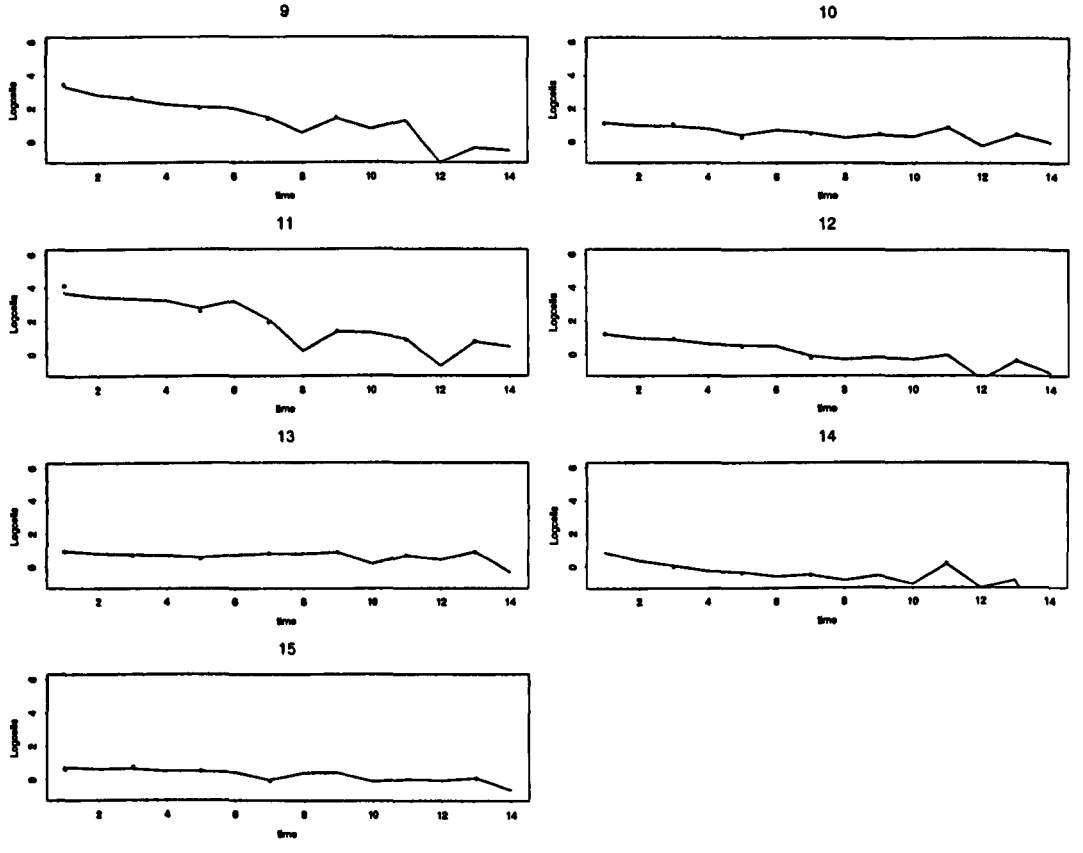
Figure 7.6: Patients 1 – 8. Dots represent observed data in log scale and solid lines estimated curves for patients 9 – 15.




very good. The kernel density estimation has been carried out using the S-Plus 'density' function that leads to a smooth estimate that could hide local features of the density. The same analysis has been performed for all the parameters of the model, leading to satisfactory results in all cases.

The expected values of the posterior distributions will be used as point estimators of the real unknown values. Substituing these point estimators in equation (7.6) together with the corresponding assumed drug concentration levels, individual predicted curves are obtained. In figures 7.6 and 7.7, observed points are represented by dots and individual predicted curves by solid lines.

Figure 7.7: Patients 9 – 15. Dots represent observed data in log scale and solid lines estimated curves for patients 9 – 15.



In general terms, the predicted individual curves in figures 7.6 and 7.7 are very encouraging. The model seem to be adequate to predict the evolution of the number of leukaemic cells in the organism of the patients. Keep in mind that the pharmacokinetic part of the model has been kept fixed as no data about drug concentration levels have been reported. Nevertheless, the model we have constructed, has proved to be able to capture the remaining variability in the model. In consequence, it seems that when the pharmacokinetic processes of all drugs involved in a particular treatment are well understood and correctly modelled, the model and the methodology proposed in this chapter will capture the the main processes involved in determining the evolution of leukaemic cell populations.

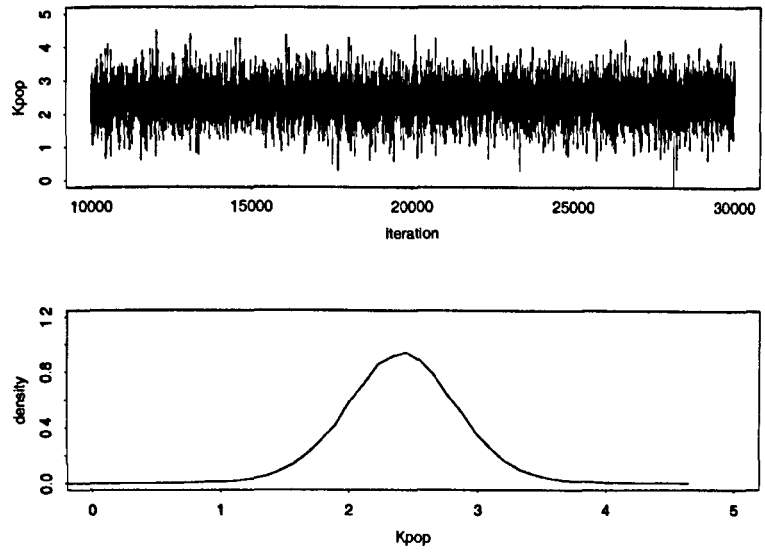
However, as we have seen in chapters 3 and 4, the pharmacokinetic processes are extremely variable both within and between individuals, and great difficoulties arise when trying to model them. In addition, when several different drugs are being administered simultaneously, complicated drug interactions might occur. We, therefore, believe that research in the pharmacokinetics of drug 'cocktails' plays a vital role in the study of leukaemia.

## 7.6 Convergence Analysis

As discussed in chapter 3, all the analysis presented in section 7.5 requires that the Gibbs sampler and the Metropolis-Hastings algorithm have effectively converged to their unique stationary or invariant distributions. Convergence diagnostics are fundamental in order to establish that the sample obtained by running each algorithm constitutes a sample from the corresponding posterior distribution.

As briefly discussed in chapter 3, much research is being made for diagnosing convergence of MCMC algorithms (see references in chapter 3). No particular

Figure 7.8: Trace and Posterior Distribution obtained for the population intercept, $K_p$. The first plot shows the evolution of the chain in the 20.000 iterations after the burn-in period. The inferior figure is the corresponding estimated posterior distribution.




method of assessing convergence is universally accepted as being superior in different scenarios. In this work, only a graphical inspection of the chains has been performed. As these seem to mix rapidly over the sample space of the posterior distributions, we have not performed any further study. As an illustration, the trace obtained for the population intercept, $K_p$, is shown in figure 7.8.

## 7.7 Conclusions

In this last chapter, the tools introduced in previous chapters have been applied to real data. Three interrelated complicated biological processes have been encompassed in a global model: pharmacokinetics, pharmacodynamics and cell kinetics. Due to the lack of pharmacokinetic data, an analysis of the

pharmacokinetic processes of the different drugs that constitute treatment has not been possible.

The Bayesian population approach has proved to be remarkably appropriate in order to capture the variability that usually cell kinetic and pharmacodynamic data show. Nevertheless, when the pharmacokinetic parameters of the drugs involved in treatment need to be estimated, the variability and the number of parameters of the model would increase dramatically.

Future work should be oriented to analysing identifiability of model parameters when all the three processes are being modelled simultaneously. Complicated pharmacokinetic models are usually needed when different drugs are being administered simultaneously. Poor pharmacokinetic estimations would lead to poor pharmacodynamic and cell kinetic estimations, and therefore, the results could not be used for dose recommendation.

# Chapter 8

# Conclusions

A better understanding of the treatment-effect-time relationship can lead to more efficient treatments based on individual characteristics and disease intensity profiles. Innumerable interrelated biologically complex processes are involved in determining the ultimate effect that drugs initiate in the organism of patients.These can be grouped under the headings ofPharmacokinetic and Pharmacodynamic processes.

In chapter 2 the main pharmacokinetic processes have been introduced and appropriate compartmental models constructed. The essential aim of these models is to explain and predict the evolution of drugs within the body by modelling the relationship between drug doses and drug concentration levels in the blood and other sites of interest. The compartmental approach conceptualises the body as a compound of different compartments and it analyses the kinetics between and within them.

Pharmacokinetic data (*i.e.*, drug concentration levels found in the blood at different time points) usually present considerable variability both within and between individuals. Due to its attempt to explain and control this variability, the population approach has gained prevalence over the individual pharma-

cokinetic approach during the last two decades. In this work, the population pharmacokinetic/pharmacodynamic approach has been taken.

The population approach has been studied in chapter 3 within the Bayesian framework that, primarily due to recent enormous advances in 'Markov chain Monte Carlo' techniques, offers flexible and simple methods to build and estimate population pharmacokinetic models. By exploiting the hierarchical structure of the models of interest we have easily built a general population pharmacokinetic model that accommodates most of the models proposed in the literature for different drugs.

The strength of the population approach has been assessed in chapter 4. A population pharmacokinetic analyses has been performed in order to advance the study of the pharmacokinetics of 'cyclosporin', a drug in routine use in organ transplantation processes. A two-compartment model has been assumed to explain the evolution of cyclosporin blood concentration levels following certain dosing histories. A fully Bayesian population model has been specified and all model parameters estimated. The results support the adequacy of the population approach to handle sparse pharmacokinetic data.

The ultimate aim of this work is to describe an approach to constructing a model that both quantitatively and qualitatively relates drug administration to consequent changes in complicated biological processes. We have concentrated our attention to the evolution of leukaemic cell populations. In order to identify the drug-specific effect in such populations, the kinetics of leukaemic cells before treatment starts need to be understood. In chapter 5 a qualitative description of the evolution of leukaemic cell populations has been followed by several models that capture the main processes that govern cell proliferation.

Relying on different assumptions, two approaches have been taken in order to design stochastic models for cell proliferation. The Multi-Type Birth and

Death process in section 5.3.3 presents the advantage of being able to identify and follow the evolution of different cell subpopulations according to their role in the cell cycle. The underlying key assumption is that the life lengths of the cells are exponentially distributed so that transitions from phase to phase occur at random.

The Multiple-Phase Birth and Death Process built in section 5.4 constitutes an alternative to exponentially distributed cell life lengths without renouncing the tractable continuous time Markov process theory. The main disadvantage of this model is that it does not allow general life length distributions either and that it cannot distinguish the evolution of subpopulations with different kinetic behaviours.

As an attempt to overcome the limitation of constrained cell life length's distributions, a Branching process admitting general life length distributions has been constructed in section 5.5. A limit result has been shown for the expected number of cells in the population and an interesting and simple approximation obtained. This will allow the prediction of the expected evolution of a cell population with knowledge only about the mean and variance of individual cell's life length distributions. Future work would include a deeper evaluation of this result.

After analysing the biological process that will be modified by the treatment, we have gone to model the effect originated by drugs as a function of drug levels found in the body. A fully Bayesian Population Pharmacokinetic-Pharmacodynamic model has been built in chapter 6. Several simulation studies have been performed in order to illustrate the potential of these models. The crucial phenomenon of drug ressistance has been acknowledged and introduced in the model.

In this work three connected processes have been modelled (*i.e.* pharma-

cokinetics, cell kinetics and pharmacodynamics) and relationships between them considered so that a global model linking treatment and the evolution of leukaemic cell populations has been proposed. The population approach has been advocated as a powerful framework to handle real data variability.

In the last chapter of this work a real data set has been presented. Data about treatment history and the number of leukaemic cells in the organism at different time points of a group of patients has been analysed. The absence of data about blood drug concentration levels has forced us to arbitrarily set individual pharmacokinetic parameters and, therefore, to consider a simplified form of the model constructed in chapter 6. Consequently, the results obtained cannot be used for prediction purposes as they rely on the assumed blood drug concentration levels.

Nevertheless, the model has proved to be able to capture the underlying main pharmacodynamic and cell kinetic processes and to generate good estimators of the parameters. In a realistic setting, where no perfect knowledge of the pharmacokinetics of each of the drugs is available, parameter identifiability would become a complicated task. Drug interactions have shown in the literature to be complicated to model. Great difficulties arise when trying to model simultaneously the pharmacokinetics of different drugs. The first step would be to get a better understanding of individual pharmacokinetic characteristics of each of the drugs that constitute treatment for different patients. Then, the interaction of the drugs should be studied and an adequate model constructed. Possible identifiability problems should then be overcome.

All this prior information would facilitate the specification and estimation of the general model. The better the understanding of each of the three processes that are considered in the general model (*i.e.*, pharmacokinetics, pharmacodynamics and cell kinetics), the better the estimation of the whole model.

We leave for future work the analysis of a data set containing treatment regimes, blood drug concentration levels and the number of leukaemic cells in the organism. The ambitious aim will be to get population and individual parameter estimates so that there is better knowledge when designing both new treatments for new patients and adjusted treatments for patients for whom estimates have already been obtained.

The complexity of the task is obvious, but any tiny improvement in the rationale behind individual treatments could make a great difference in the fight against leukaemia.

# Appendix A

# Individual Pharmacokinetic Parameters' Posterior Distributions

Table A.1: Summary: Marginal Posterior Distributions for Individual Pharmacokinetic Parameters.

| Indiv | ln(Param) | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|---|
| Patient 1 | $\mu_1 = C_L$ | -2.016 | 0.06841 | 0.002177 | -2.147 | -2.014 | -1.895 |
| | $\mu_2 = Q$ | -1.12 | 0.2302 | 0.00451 | -1.542 | -1.128 | -0.6311 |
| | $\mu_3 = V_1$ | 3.261 | 0.2731 | 0.00737 | 2.691 | 3.269 | 3.77 |
| | $\mu_4 = V_2$ | 4.822 | 0.4782 | 0.01651 | 4.061 | 4.758 | 5.872 |
| Patient 2 | $\mu_1 = C_L$ | -1.395 | 0.1225 | 0.006332 | -1.719 | -1.375 | -1.23 |
| | $\mu_2 = Q$ | -0.7618 | 0.2138 | 0.004444 | -1.175 | -0.7631 | -0.3375 |
| | $\mu_3 = V_1$ | 3.786 | 0.2699 | 0.007114 | 3.249 | 3.783 | 4.313 |
| | $\mu_4 = V_2$ | 5.794 | 0.6109 | 0.03059 | 4.787 | 5.726 | 7.1037 |
| Patient 3 | $\mu_1 = C_L$ | -0.852 | 0.2631 | 0.01915 | -1.674 | -0.7599 | -0.6132 |
| | $\mu_2 = Q$ | 0.02406 | 0.6263 | 0.02811 | -1.042 | -0.02739 | 1.803 |
| | $\mu_3 = V_1$ | 5.422 | 0.2186 | 0.005613 | 4.982 | 5.432 | 5.82 |
| | $\mu_4 = V_2$ | 6.563 | 1.026 | 0.06449 | 4.361 | 6.659 | 8.285 |

Table A.2: Summary: Marginal Posterior Distributions for Individual Pharmacokinetic Parameters.

| Indiv | ln(Param) | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|---|
| Patient 4 | $\mu_1 = C_L$ | -1.543 | 0.06916 | 0.002078 | -1.672 | -1.539 | -1.425 |
| | $\mu_2 = Q$ | -0.5803 | 0.2675 | 0.006 | -1.063 | -0.5919 | -0.001044 |
| | $\mu_3 = V_1$ | 3.735 | 0.3365 | 0.00915 | 3.011 | 3.756 | 4.328 |
| | $\mu_4 = V_2$ | 5.118 | 0.5179 | 0.01946 | 4.34 | 5.029 | 6.354 |
| Patient 5 | $\mu_1 = C_L$ | -1.91 | 0.08005 | 0.003757 | -2.076 | -1.904 | -1.785 |
| | $\mu_2 = Q$ | -1.734 | 0.2376 | 0.005676 | -2.156 | -1.742 | -1.227 |
| | $\mu_3 = V_1$ | 2.829 | 0.1959 | 0.005447 | 2.41 | 2.837 | 3.181 |
| | $\mu_4 = V_2$ | 4.465 | 0.672 | 0.03025 | 3.506 | 4.333 | 6.089 |
| Patient 6 | $\mu_1 = C_L$ | -1.756 | 0.06678 | 0.001906 | -1.882 | -1.754 | -1.637 |
| | $\mu_2 = Q$ | -0.6592 | 0.3022 | 0.008132 | -1.195 | -0.6816 | -0.008874 |
| | $\mu_3 = V_1$ | 3.457 | 0.4047 | 0.01313 | 2.577 | 3.484 | 4.17 |
| | $\mu_4 = V_2$ | 4.789 | 0.4751 | 0.01535 | 4.098 | 4.707 | 5.949 |
| Patient 7 | $\mu_1 = C_L$ | -1.836 | 0.07362 | 0.002604 | -1.995 | -1.831 | -1.714 |
| | $\mu_2 = Q$ | -0.7082 | 0.2933 | 0.007405 | -1.215 | -0.7274 | -0.06224 |
| | $\mu_3 = V_1$ | 3.568 | 0.329 | 0.00975 | 2.873 | 3.597 | 4.154 |
| | $\mu_4 = V_2$ | 4.974 | 0.5306 | 0.01986 | 4.15 | 4.886 | 6.219 |
| Patient 8 | $\mu_1 = C_L$ | -2.677 | 0.06043 | 0.00192 | -2.808 | -2.673 | -2.573 |
| | $\mu_2 = Q$ | -1.972 | 0.245 | 0.005005 | -2.428 | -1.974 | -1.459 |
| | $\mu_3 = V_1$ | 2.551 | 0.2194 | 0.005212 | 2.105 | 2.558 | 2.957 |
| | $\mu_4 = V_2$ | 4.1 | 0.5261 | 0.01775 | 3.207 | 4.068 | 5.19 |

Table A.3: Summary: Marginal Posterior Distributions for Individual Pharmacokinetic Parameters.

| Indiv | ln(Param) | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|---|
| Patient 9 | $\mu_1 = C_L$ | -1.195 | 0.08571 | 0.00385 | -1.391 | -1.189 | -1.048 |
| | $\mu_2 = Q$ | -1.014 | 0.2319 | 0.005828 | -1.47 | -1.015 | -0.5414 |
| | $\mu_3 = V_1$ | 3.546 | 0.3559 | 0.0151 | 2.763 | 3.58 | 4.148 |
| | $\mu_4 = V_2$ | 5.189 | 0.7328 | 0.037 | 4.147 | 5.038 | 6.919 |
| Patient 10 | $\mu_1 = C_L$ | -1.788 | 0.06557 | 0.001649 | -1.927 | -1.785 | -1.668 |
| | $\mu_2 = Q$ | -0.8363 | 0.2805 | 0.006445 | -1.347 | -0.8488 | -0.234 |
| | $\mu_3 = V_1$ | 3.503 | 0.3221 | 0.008822 | 2.791 | 3.53 | 4.068 |
| | $\mu_4 = V_2$ | 4.977 | 0.4889 | 0.01616 | 4.173 | 4.921 | 6.082 |
| Patient 11 | $\mu_1 = C_L$ | -2.032 | 0.07042 | 0.002452 | -2.177 | -2.027 | -1.915 |
| | $\mu_2 = Q$ | -0.7425 | 0.4222 | 0.01302 | -1.451 | 0.784 | 0.2196 |
| | $\mu_3 = V_1$ | 3.613 | 0.3577 | 0.01183 | 2.807 | 3.647 | 4.223 |
| | $\mu_4 = V_2$ | 4.752 | 0.5482 | 0.02086 | 3.829 | 4.692 | 5.951 |
| Patient 12 | $\mu_1 = C_L$ | -1.944 | 0.1035 | 0.004719 | -2.187 | -1.934 | -1.77 |
| | $\mu_2 = Q$ | -2.57 | 0.239 | 0.008542 | -3.047 | -2.566 | -2.105 |
| | $\mu_3 = V_1$ | 2.387 | 0.204 | 0.008329 | 1.942 | 2.401 | 2.765 |
| | $\mu_4 = V_2$ | 4.565 | 0.9825 | 0.05167 | 2.702 | 4.586 | 6.353 |
| Patient 13 | $\mu_1 = C_L$ | -2.09 | 0.05698 | 0.001447 | -2.206 | -2.088 | -1.98 |
| | $\mu_2 = Q$ | -0.287 | 0.7623 | 0.03913 | -1.879 | -0.2751 | 1.184 |
| | $\mu_3 = V_1$ | 2.867 | 0.4223 | 0.02366 | 1.822 | 2.937 | 3.496 |
| | $\mu_4 = V_2$ | 2.724 | 0.5099 | 0.02712 | 1.627 | 2.777 | 3.392 |

Table A.4: Summary: Marginal Posterior Distributions for Individual Pharmacokinetic Parameters.

| Indiv | ln(Param) | Mean | sd | MC error | 2.5% | median | 97.5% |
|---|---|---|---|---|---|---|---|
| Patient 14 | $\mu_1 = C_L$ | -1.703 | 0.06112 | 0.001784 | -1.829 | -1.702 | -1.589 |
| | $\mu_2 = Q$ | -1.13 | 0.3043 | 0.009855 | -1.672 | -1.151 | -0.4746 |
| | $\mu_3 = V_1$ | 3.182 | 0.3575 | 0.01442 | 2.404 | 3.21 | 3.81 |
| | $\mu_4 = V_2$ | 4.552 | 0.5923 | 0.02563 | 3.791 | 4.391 | 5.992 |
| Patient 15 | $\mu_1 = C_L$ | -1.217 | 0.133 | 0.009131 | -1.613 | -1.196 | -1.06 |
| | $\mu_2 = Q$ | -0.9437 | 0.222 | 0.006044 | -1.357 | -0.9536 | -0.4901 |
| | $\mu_3 = V_1$ | 3.809 | 0.3015 | 0.01211 | 3.157 | 3.818 | 4.356 |
| | $\mu_4 = V_2$ | 5.518 | 0.7607 | 0.04386 | 4.384 | 5.389 | 7.521 |
| Patient 16 | $\mu_1 = C_L$ | -2.343 | 0.0748 | 0.002941 | -2.492 | -2.343 | -2.201 |
| | $\mu_2 = Q$ | -2.154 | 0.2221 | 0.007476 | -2.587 | -2.158 | -1.702 |
| | $\mu_3 = V_1$ | 1.801 | 0.3084 | 0.01313 | 1.215 | 1.8 | 2.405 |
| | $\mu_4 = V_2$ | 3.872 | 0.511 | 0.023 | 3.135 | 3.775 | 5.152 |
| Patient 17 | $\mu_1 = C_L$ | -1.778 | 0.06267 | 0.001792 | -1.915 | -1.775 | -1.666 |
| | $\mu_2 = Q$ | -1.05 | 0.204 | 0.00422 | -1.438 | -1.054 | -0.6325 |
| | $\mu_3 = V_1$ | 3.146 | 0.2847 | 0.007367 | 2.559 | 3.156 | 3.682 |
| | $\mu_4 = V_2$ | 5.202 | 0.4479 | 0.01638 | 4.416 | 5.164 | 6.169 |

[1] Andersson, J.M., Kin, S.W., Kopecek, J. and Knutson, K. (1994). Proceedings of the sixth symposium on recent advances in drug delivery systems. *Journal of Controlled Release*, **28**, pp.1.

[2] Athreya, K.B. and Ney, P. *Branching Processes.* Springer-Verlag, Berlin, 1973.

[3] Artemov, D., Bhujwalla, Z.M., Pilatus, U. and Glicson, J.D. (1998). Two-compartment model for determination of glycolytic rates of solid tumors by in vivo C NMR spectroscopy. *NMR in Biomedicine*, **11**, Issue 8, pp. 395-404.

[4] Bailey, N.T.J. (1967) *The elements of stochastic processes with applications to the natural sciences.* 3*th* ed. New York: Wiley.

[5] Bellman, R. and Harris, T.E. (1952). On age-dependent binary branching processes. *Annals of Mathematics*, **55**, pp. 280-295.

[6] Benet, L.Z. and Galeazzi, R.L. (1979). Noncompartmental determination of the steady-state volume of distribution. *Journal of Pharmaceutical Science*, **48**, pp. 1071.

[7] Bennet, J.E., Racine-Poon, A. and Wakefield, J.C. MCMC for nonlinear hierarchical models. In *Markov Chain Monte Carlos in Practice.* Eds. Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. pp. 339-357. Chapman and Hall, London, 1995.

[8] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis* (second edition). New York: Springer-Verlag.

[9] Berkey, C.S. (1982). Bayesian approach for a nonlinear growth model. *Biometrics*, **38**, pp. 953-961.

[10] Bernardo, J.M. and Smith, A.F.M. (1994). *Bayesian Theory.* New York: Wiley.

[11] Besag, J., Green, P.J., Higdon, D. and Mengersen, K. (1995). Bayesian computation and stochastic systems (with discussion). *Journal of the Royal Statistical Society, Series B*, **10**, pp. 3-66.

[12] Biggins, J.D., Cohn, H. and Nerman, O. (1999). Multi-type branching in varying environment. *Stochastic Processes and their Applications*, **83**, pp. 357-400.

[13] Breant, V., Charpiat, B., Sab, J.M., Maire, P, and Jeliffe, R.W. (1996). How many patients and blood levels are necessary for population pharmacokinetic analysis? A study of a one-compartment model applied to cyclosporine. *European Journal of Clinical Pharmacology*, **51**, Issue 3-4, pp. 283-288.

[14] Brooks, S. and Roberts, G. (1995). *A review of diagnostic methods for Markov chain Monte Carlo.* Technical report, Dept of Pure Mathematics and Mathematical Statistics, University of Cambridge.

[15] Brooks, S.P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, **7**, pp. 435-355.

[16] Brooks, S.P. and Roberts, G.O. (1998). Convergence assessment techniques for Markov chain Monte Carlos. *Statistics and Computing*, **8**, pp. 319-335.

[17] Burguess, A.W. and Nicola, N.A. *Growth factors and Stem Cells.* Academic Press, New York, 1983.

[18] Catling, S.N., Abkowitz, J.L. and Guttorp, P. (2001). Statistical inference in a two-compartment model for hematopoiesis. *Biometrics*, **57**,

Issue 2, pp. 546-553.

[19] Chappell, M.J. (1995). Structural identifiability and indistinguishability of certain two-compartment models incorporating nonlinear efflux from the peripheral compartment. *Mathematical Biosciences*, **121**, Issue 1, pp. 61-81.

[20] Charpiat, B., Falconi, I., Breant, V., Jellife, R.W., Sab, J.M., Ducerf, C., Fourcade, N., Thomasson, A. and Baulieux, J. (1998). A population pharmacokinetic model of cyclosporine in the early post-operative phase in patients with liver transplants, and its predictive performance with Bayesian fitting. *Therapeutic Drug Monitoring*, **20**, pp. 158-164.

[21] Chiang, C.L. (1980). *An Introduction to Stochastic Processes and Their Applications.* Krieger, R.E, Melbourne, FL.

[22] Chieck, S., Shortle, J, van Gelder, P. and Mendel, M.B. (1996). A model for the frequency of extreme river levels based on river dynamics. *Structural Safety*, **18**, pp. 261-276.

[23] Chinsan, L. (1998). An Integral characterization problem in an age-dependent cancer model. *Statistics & Probability Letters*, **40**, pp. 121-126.

[24] Christensen, K.J. and Shonkweiller, R. (1978). Lag time in microbe growth as an age-dependent branching process with two phase types. *Bulletin of Mathematical Biology*, 40, pp. 347-358.

[25] Collins, J.M., Grieshaber, C.K. and Chabner, B.A. (1990) Pharmacologically guided phase I clinical trials based upon preclinical drug development. *Journal of the National Cancer Institute*, **82**, pp. 1321-1326.

[26] Cowan, R. (1985). Branching process results in terms of moments of the generation-time distribution. *Biometrics*, **41**, pp. 681-689.

[27] Cowan, R. and Morris, V.B. (1986). Cell population dynamics during the differentiative phase of tissue development. *Journal of Theoretical Biology*, **122**, pp. 205-224.

[28] Cowles, M.K. and Carlin, B.P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, **91**, pp. 883-904.

[29] Dayneka, N.L., Garg, V. and Jusco, W.J. (1993). Comparison of four basic models of indirect pharmacodynamic responses. *Journal of Pharmacokinetic Biopharmacology*, **21**, pp. 125.

[30] Creasy, G.W. and Jaffe, M.E. (1991). Endocrine-reproductive pulsatile delivery systems. *Advanced Drug Delivery Reviews*, **6**, pp. 917.

[31] Davis, P.J. and P. Rabinowitz (1994). *Methods of Numerical Integration.* Orlando: Academic Press. (second edition).

[32] DuMouchel, W. and Waternaux, C. (1992). Discussion on hierarchical models for combining information and for meta-analyses (by C.N. Morris and S.L. Normand). In *Bayesian Statistics 4* (eds J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F. M. Smith), pp. 338-341. Oxford: Oxford University Press.

[33] Ebling, W.F., Danhof, M. and Stanski, D.R. (1991). Pharmacodynamic characterization of the electroencephalographic effects of thiopental in rats, *Journal of Pharmcokinetics and Biopharmacology. 19*, pp. 123.

[34] Edgren, B., Karna, P., Sciamanna, D. and Dolanski, E. (1984). Gentamicin dosing in the newborn. Use of a one-compartment open pharmacokinetic model to individualize dosing. *Dev. Pharmacol. Ther.*, **7**, Issue 4, pp. 263-272.

[35] Elton, C.S. *The Ecology of Invasions by Animals and Plants*, Methuen, London, 1958.

[36] Fahr, A. (1993). Cyclosporine Clinical Pharmacokinetics. *Clinical Pharmacokinetics*, **24**, pp. 472-495.

[37] Garthwaite, P.H. and Dickey, J.M. (1992). Elicitation of prior distributions for variable selection problems in regression. *Annals of Statistics*, **20**, pp. 1697-1719.

[38] Gelfand, A.E., Hills, S.E., Racine-Poon, A. and Smith, A.F.M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Am. Statist. Ass.* **85**, pp. 972-985.

[39] Gelfand, A.E. and Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *J. Am. Statist. Ass.* **85**, pp. 398-409.

[40] Gelman, A. and Rubin, D.M. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 4, pp. 457-472.

[41] Gelman, A., Roberts, G.O. and Gilks, W.R. (1996). Efficient Metropolis jumping rules, in *Bayesian Statistics 5*, eds. J.M. Bernardo *et al*), Oxford University Press, Oxford, pp. 599-607.

[42] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith (eds), *Procedings of the Fourth Valencia International Conference on Bayesian Statistics*, pp. 169-193. New York: Oxford University Press.

[43] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattn. Anal. Mach. Intel.* **6**, pp. 721-741.

[44] Gero, A. (1971). Intimate study of drug action III: Mechanisms of molecular drug action. In *Drill's Pharmacology in Medicine.* McGraw-Hill.

[45] Gibaldi, M. and Perrier, D. *Pharmacokinetics, 2nd ed. Revised and Expanded.* Marcel Dekker Inc., New York, 1982.

[46] Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (eds.) (1996). *Markov chain Monte Carlo in practice*, Chapman and Hall, London.

[47] Gillespie, W.R. (1991). Noncompartmental versus compartmental modelling in clinical pharmacokinetics. *Clinical Pharmacokinetics, 20*, pp. 253.

[48] Harris, T.T. *The Theory of Branching Processes.* Springer-Verlag. Berlin (1963).

[49] Hastings, W.K. (1970). Monte Carlos sampling methods using markov chains and their applications. *Biometrica*, **57**, pp. 97-109.

[50] Heidelberger, P. and Welch, P. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, **31**, pp. 1109-1144.

[51] Hengeveld, R. *Dynamics of Biological Invasions.* Chapman and Hall, London, 1989.

[52] Hill, A.B. (1910). The possible effets of the aggregation of the molecules of haemoglobin on its dissociation curves. *Journal of Physiology*, **40**, pp. iv-vii.

[53] Hoke, J.F. and Ravis, W.R. (1992). Evaluation of methods for estimating the rate constant of a one-compartment absorption model when absorption and elimination rate constants are equal. *Journal of Pharmaceutical Sciences*, **81**, Issue 4, pp. 401-402.

[54] Howard, A. and Pelc, S.R. (1970). Synthesis of desoxiribonucleic acid in normal and irradiated cells and its relation to chromosome breakage. *Heredity Suppl.*, **6**, pp. 261-273.

[55] Jagers, P. *Branching Processes with Biological Applications.* Wiley, London, 1975.

[56] Jagers, P. and Klebaner, F.C. (2000). Population-size-dependent and age-dependent branching processes. *Stochastic Processes and their*

*Applications*, **87**, pp. 235-254.

[57] Jaiswal, N.K., Sangeeta, Y. and Gaur, S.C. (1997). Stochastic modelling of combat with reinforcement. *European Journal of Operational Research*, **100**, pp. 225-235.

[58] Jenkinson, D.H., Barnard, E.A., Hoyer, D. *et al.* International union of pharmacology committee on receptor nomenclature and drug classification. IX. Recommendations on terms and symbols in quantitative pharmacology. *Journal of Pharmacological Science*, **47**, pp. 225.

[59] Jennison, C. (1993). Discussion of 'Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods', by Smith and Roberts, *Journal of the Royal Statistical Society, Series B*, **55**, pp. 54-56.

[60] Jensen, S.A. and Dalhoff, K.P. (2000). Therapeutic monitoring by blood concentrations with the focus on cyclosporin A. *Ugeskrift for Laeger*, **163**, pp. 2009-2012.

[61] Jusco, W.J. (1971). Pharmacodynamics of chemotherapeutic effects: Dose-time-response relationships for phase-non-specific agents. *Journal of Pharm. Sciences*, **60**, pp. 892.

[62] Jusco, W.J. (1990). Corticosteroid pharmacodynamics: models for a broad array of receptor-mediated pharmacologic effects. *Journal of Clinical Pharmacology*, **30**, pp. 303.

[63] Jusco, W.J. Guidelines for collection and analysis of pharmacokinetic data, In Applied Pharmacokinetics: Principles of Therapeutic Drug Monitoring, 3*rd* ed. Evans, W.E., Schentag, J.J. and Jusco, W.J. eds. *Applied Therapeutics*, Spoknae, W.A., 1982, 1992.

[64] Jusco, W.J. and Ko, H.C. (1994). Physiological indirect response models characterize diverse types of pharmacodynamic effects. *Clinical Pharmacology and Therapeutics*, **56**, pp. 406.

[65] Kenakin, T. (1993). *Pharmacologic analysis of drug-receptor interaction.* Raven Press, New York.

[66] Kumperscak, M. and Kozjek, F. (1998). Non-compartment pharmacokinetic analysis- Physiological models. *Farmacewtski Vestnik*, **49**, Issue 2, pp.167-177.

[67] Lassen, N.A. and Perl, W. (1979). *Trace kinetic methods in medical physiology.* Raven Press, New York.

[68] Lauritzen, S.L., Dawid, A.P., Larsen, B.N. and Leimer, H.G. (1990). Independence properties of directed Markov fields. *Networks* **20**, pp. 491-505.

[69] Lauritzen, S.L (1996). *Graphical Models.* Oxford: Clarendon.

[70] Lee, C.J., Hsu, S.T. and Hu, S.C. (1989). A one-compartment single-pore model for extracorporeal hemoperfusion. *Comput. Biol. Med*, **19**, Issue 2, pp. 83-94.

[71] Lefevre, C. (1981). Optimal control of a birth and death epidemic process. *Operations Research*, **29**, pp. 971-982.

[72] Levy, G. (1966). Kinetics of pharmacological effects. *Clinical Pharmacology and Therapeutics*, **7**, pp 362.

[73] Levy, R.H. (1982). Time-dependent pharmacokinetics. In *Pharmacokinetics: Theory and Methodology.* Edited by Rowland, M. and Tucker, G. Pergamon Press.

[74] Levy, G. (1994). Mechanism-based pharmacodynamic modeling. *Clinical Pharmacology and Therapeutics*, **56**, pp. 356.

[75] Lewitt, M.D. and Lewitt, D.G. (1998). Use of a two-compartment model to assess the pharmacokinetics of human ethanol metabolism. *Alcoholism, Clinical and Experimental Research*, **22**, Issue 8, pp. 1680-1688.

[76] Lindley, D.V. (1970). The estimation of many parameters. In *Foundation of statistical inference.* Holt, Rinehart and Winston, Toronto.

[77] Lindstrom, M.J. and Bates, D.M. (1990) Nonlinear mixed effects models for repeated measures data . *Biometrics*, **46**, pp. 673-687.

[78] Mathews, J.C. *Fundamentals of Receptor, enzyme and transport kinetics.* CRC *Press.*, Bocan Raton (1993).

[79] Matucci, M., Longo, G., Morfini, M., Vannini, S., Messori, A., Donati-Cori, G. and Tendi, E. (1983). Pharmacokinetics of factor VIII: C. Appropriateness of the one-compartment model for estimating clearance. *Farmaco (Societa Chimica Italiana: 1989)*, **38**, Issue 8, pp. 306-311.

[80] McCarley, K.D. and Bunge, A.L. (2000). Physiologically relevant two-compartment pharmacokinetic models for skin. *Journal of Pharmaceutical Sciences, 89*, Issue 9, pp. 1212-1235.

[81] Mengersen, K.L. and Tweedie, R.L. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Annals of Statistics*, **24**, pp. 101-121.

[82] Merkhofer, N.W. (1987). Quantifying judgemental uncertainty: methodology, experiences and insights. *IEEE Trans. Systems, Science and Cybernetics*,**17**, pp.741-752.

[83] Metropolis, N., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, **21**, pp. 1087-1092.

[84] Metz, H.A., Dienseke, H., de Jonge, G. and Putters, F.A. (1983). Continuous-time Markov chains as models for animal behaviour. *Bulletin of Mathematical Biology*, **45**, pp. 643-658.

[85] Meyn, S.P. and Tweedie, R.L. (1993). *Markov chains and stochastic stability.* Springer-Verlag, New York.

[86] Murakami, H., Yamane, I., Barnes, D.W. *et al.* (eds): *Growth and Differentiation of Cells in Defined Environment.* New York, Springer-Verlag, 1985.

[87] Murray, A. and Wood, A. (1993) *The cell cycle.* Oxford University Press.

[88] Nakashima, E. and Benet L.Z. (1989). An integrated approach to pharmacokinetic analysis for linear mammillary systems in which input and exit may occur in/from any compartment. *Jounal of Pharmacokinetic and Biopharmacology*, **17**, pp. 673.

[89] Norberg, R. (1995). Differential equations for moments of present values in life insurance. *Insurance: Mathematics and Economics*, **17**, pp. 171-180.

[90] Norman, J. (1992). One-compartment kinetics. *British Journal of Anaesthesia*, **69**, Issue 4, pp. 387-396.

[91] Novick, M.R., Jackson, P.H., Thayer, D.T. and Cole, N.S. (1972) Estimating multiple regression in $m$-groups; a cross-validation study. *Br. J. Math. Statist. Psychol.*, **25**, pp. 33-50.

[92] Parke, J. and Charles, B.G. (1998). NONMEM population pharmacokinetic modeling of orally administered cyclosporine from routine drug monitoring data after heart transplantation. *Therapeutic Drug Monitoring*, **20**, pp. 284-293.

[93] Peskun, P.H. (1973). Optimum Monte-Carlo Sampling using Markov Chains. *Biometrica*, **60**, pp. 607-612.

[94] PKBUGS, Version 1.1. Imperial College of Science, Technology and Medicine, 1999.

[95] Polson, N.G. (1994). Convergence of Markov chain Monte Carlo algorithms. In J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (eds), *Proceedings of the Fifth Valencia International Conference on Bayesian Statistics.* New York: Oxford University Press.

[96] Purves, R.D. (1993). Anomalous parameter estimates in the one-compartment model with first-order absorption. *Journal of Pharmacy and Pharmacology*, **45**, Issue 10, pp. 934-936.

[97] Racine-Poon, A. and Wakefield, J. (1996). Statistical methods for population pharmacokinetic modelling. *Statistical Methods in Medical Research*, **7**, Issue 1, pp. 63-84.

[98] Raftery, A.E. and Lewis, S.M. (1992). How many iterations in the Gibbs sampler? In J.M. Bernardo , J.O. Berger, A.P. Dawid and A.F.M. Smith (eds), *Proceedings of the Fifth Valencia International Conference on Bayesian Statistics*, pp. 763-774. New York: Oxford University Press.

[99] Renshaw, E. *Modeling Biological Populations in Space and Time.* Cambridge University Press, New York, 1991.

[100] Renshaw, E. (2000). Applying the saddlepoint approximation to bivariate stochastic processes. *Mathematical Biosciences*, **168**, pp. 57-75.

[101] Rescigno, A. (2001). The rise and fall of compartmental analysis. *Pharmacological Research*, *44*, No. 4, pp. 337-342.

[102] Ritter, C. and Tanner, M.A. (1992). The Gibbs stopper and the griddy Gibbs sampler. *Journal of the American Statistical Association* **87**, pp. 861-868.

[103] Roberts, G.O. and Smith, A.F.M. (1994). Some convergence theory for Markov chain Monte Carlo. *Stochastic Processes and Applications* **49**, pp. 207-216.

[104] Roberts, G.O. and Tweedie, R.L. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometric*, **83**, pp. 95-100.

[105] Roberts, G.O. and Sahu, K. (1997). Updating schemes, correlation structure, blocking and parameterization for the Gibbs sampler. *J.R.*

*Statist. Soc. B* **59**, pp. 291-317.c

[106] Robinson, I. (1991). Chronopharmacology of growth hormone and related peptides. *Advanced Drug Delivery Reviews*, **6**, pp.57

[107] Ross, E.M. *Pharmacodynamics: Mechanisms of drug action and the relationship between drug concentration and effect.* In Goodman Gilman's *The Pharmacological basis of therapeutics.* $10^{th}$ edition. Pergamom Press, New York (1996).

[108] c Ross, S.M. *Stochastic Processes*, *2nd* ed. John Wiley & Sons, Inc. New York, 1996.

[109] Rowland, M. and Tozer, T. *Clinical Pharmacokinetics: Concepts and Applications*, *3rd* ed. Lea and Febiger. New York, 1995.

[110] Rui, J.Z., Zhuo, H.Y., Jiang, G.H. and Chen, G. (1995). Evaluation of population pharmacokinetics of cyclosporin A in renal transplantation patients with NONMEM. *Acta Pharmaceutica Sinica*, **30**, pp. 241-247.

[111] Sabot, C., Debord, J., Roullet, B., Marquet, P., Merle, L. and Lachatre, G. (1995). Comparison of a 2- and 3-compartment models for the Bayesian estimation of methotrexate pharmacokinetics. *International Journal of Clinical Pharmacology and Therapeutics*, **33**, Issue3, pp.164-169.

[112] Scott, J.C. and Stanski, D.R. (1987) Decreased fentanyl and alfentanil dose requirements with age. A simultaneous pharmacokinetic and pharmacodynamic evaluation . *Journal of Pharmacol Exp Ter*, **240**, pp. 159-172.

[113] Sevast'yanov, B.A. (1964). Age-cdependent branching processes, *Theor. Prob. Appl.*, **9**, pp. 521-527.

[114] Sheiner, L.B., Stanski, D.R., Vozeh, S., Miller, R.D. and Ham, (1979). J. Simultaneous modeling of pharmacokinetics and pharmacodynamics:

Application to d-tubocurarine. em Clinical Pharmacology and Therapeutics, **25**, pp. 358.

[115] Smye, S.W. and Will, E.J. (1995). A mathematical analysis of a two-compartment model of urea kinetics. *Physics in Medicine and Biology*, **40**, Issue 12, pp. 2005-2014. c

[116] Sheiner, L.B. (1991). The intellectual health of clinical drug evaluation. *Clinical Pharmacology Therapy*, **50**, pp. 4-9.

[117] Spiegelhalter, D.J., Best, N.G., Gilks, W.R. and Inskip, H. (1995). Hepatitis B: a case study in MCMC methods. In *Markov Chain Monte Carlo in Practice* (eds W.R. Gilks, S. Richardson and D.J. Spiegelhalter), pp. 21-44. London:Chapman and Hall.

[118] Stanski, D.R. and Maitre, P.O. (1990) Population pharmacokinetics and pharmacodynamics of thiopental: the effect of age revisited. *Anesthesiology*, **72**, pp. 412-422.

[119] Steel, G.G. (1968). Cell loss from experimental tumours. *Cell Tissue Kinet.*. **1**, pp. 193-207.

[120] Steel, G.G. (1977). *Growth Kinetics of Tumors*, Clarendon Press, Oxford.

[121] Steimer, J.L., Vozeh, S., Racine-Poon, A. Holford, N.G. and O'Neil, R. 'The population approach: rationale, methods, and applications in clinical pharmacology and drug development', in *Handbook of Experimental Pharmacology*, **110**, *Pharmacokinetics of Drugs*, Springer-Verlag, Heidelberg, 1994.

[122] Swamy, P.A.V.B. (1970) Efficient inference in a random coefficient regression model. *Econometrica*, **38**, pp. 311-323.

[123] Tierney, L. (1994). Markov chains for exploring posterior distributions. *Annals of Statistics*, **22**, pp. 1701-1762.

[124] Tierney, L. (1996). Introduction to general state-space Markov chain theory. Chapter 4 *pp. 59-74) of *Practical Markov chain Monte Carlo*,

W.R. Gilks, S. Richardson and D.J. Spiegelhalter, eds. Chapman and Hall, London.

[125] Tingle, M.D. and Park, B.K. (1993). The use of a three compartment in vitro model to investigate the role of hepatic drug metabolism in drug-induced blood dyscrasias. *British Journal of Clinical Pharmacology*, **36**, Issue 1, pp.31-38.

[126] Tsoukias, N.M. and George, S.C. (1998). A two-compartment model of pulmonary nitric oxide exchange dynamics. *Journal of Applied Physiology*, **85**, Issue 2, pp. 653-666.

[127] Unadkat, J.D., Sheiner, L.B., Hennis, P.J., Cronnelly, R., Miller, R.D. and Sharma, M. (1986) An integrated model for the interaction of muscle relaxants with their antagonists. *Journal of Applied Physiology*, **61**, pp. 1593-1598.

[128] Verotta, D., Kitts, J., Rodriguez, R., Coldwell, J., Miller, R.D. and Sheiner, L.B. (1991) Reversal of neuromuscular blockade in humans by neostigmine and edrophonium: a mathematical model. *Journal of Pharmacokinetic Biopharmacology*, **19**, pp. 713-729.

[129] Vincent, P.C. (1990). In *William Damieshek and Frederick Gunz's Leukemia.* 5em th ed. Henderson, E.S. and Lister, T.A. (ed). Philadelphia London W.B. Saunders.

[130] Vozeh, S., Steimer, J.L., Rowland, M., Morselli, P., Mentre, F. and Balant, L.P. (1996). The use of population pharmacokinetics in drug development. *Clinical Pharmacokinetics*, **30**, Issue 2, pp. 81-93.

[131] Wagner, J.G. (1968). Kinetics of pharmacological response I. Proposed relationships between response and drug concentration in the intact animal and man. *Journal of Theorethical Biology*, **20**, pp. 171.

[132] Wakefield, J.C., Smith, A.F.M., Racine-Poon, A. and Gelfand, A.E. (1994). Bayesian analysis of linear and nonlinear population mod-

els using the Gibbs sampler. *Applied Statistics*, **43**, pp. 201-221.

[133] Wakefield, J.C. and Racine-Poon, A. (1995). An application of bayesian population pharmacokinetic/pharmacodynamic models to dose recommendation. *Statistics in Medicine*, **14**, pp. 971-986.

[134] Watary, N. and Benet, L. (1989). Determination of mean input time, mean residence time, and steady-state volume of distribution with multiple drug imputs. *Journal of Pharmacokinetics and Biopharmacy*, **17**, pp. 941.

[135] Wehrly, T.E., Matis, J.H. and Otis, G.W. Approximating multivariate distributions in stochastic models of insect population dynamics, in *Multivariate Environmental Statistics*, Patil, G.P. and Rao, eds., pp. 573-596. Elsevier, New York, 1993.

[136] West, M. and Crosse, J. (1992). Modelling probabilistic agent opinion. *J. Roy. Statist. Soc. B*, **54**, pp. 285-299.

[137] White, R.A. and Terry, N.H.A. (2000). Cell Kinetics: Mathematical Models and Experimental Bases. *Mathematical and Computer Modelling*, **32**, pp. 113-124.

[138] Whittaker, J. (1990). *Graphical Models in Applied Multivariate Analysis.* Chichester: Wiley.

[139] Whitting, B., Holford, N.H.G. and Sheiner, L.B. (1980). Quantitative analysis of the disopyramide concentration-effect relationshipt. *British Journal of Clinical Pharmacology*, **9**, pp. 67.

[140] Wright, N. and Alison, M. i*The Biologicy of Epithelial Cell Populations.* Clarendon Press, Oxford, 1984

[141] Wu, G. and Furlanut, M. (1998). Prediction of serum vancomycin concentrations using one-, two- and three-compartment models with implemented population pharmacokinetic parameters and with the Bayesian method. *The Journal of Pharmacy and Pharmacology*, **50**,

Issue 8, pp. 851-856.

[142] Wu, G. (2000). Sensitivity analysis of pharmacokinetic parameters in one-compartment models. *Pharmacological Research*, **41**, Issue 4, pp. 445-453.

[143] Yashiro, M., Muso, E., Matsushima, M., Nagura, R., Sawanishi, K. and Sasayama, S. (1994). Two-compartment model of cholesterol kinetics for establishment of treatment strategy of LDL apheresis in nephrotic hypercholesterolemia. *Blood Purification*, **12**, Issue 6, pp. 317-326.

[144] Yukawa, E., Suematu, F., Yukawa, M., Minemoto, M. Ohdo, S., Higuchi, S., Goto, Y. and Aoyama, T. (2001). Population pharmacokinetics of digoxin in japanese patients: A 2-compartment pharmacokinetic model. *Clinical Pharmacokinetics*, **40**, Issue 10, pp. 773-781.

[145] Yule, U. (1924). A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis. *F.R.S., Phil. Trans*, **B 213**, pp. 21-87.

[146] Zellner, A. and Min, C. (1995). Gibbs sampler convergence criteria. *Journal of the American Statistical Association* **90**, pp. 921-927.

[147] Zorin, A., Mayer-Proschel, M., Noble, M and Yakovlev, A.Y. (2000). Estimation problems associated with stochastic modeling of proliferation and differentiation of O-2A progenitor cells in vitro. *Mathematical Biosciences*, **167**, pp. 109-121.